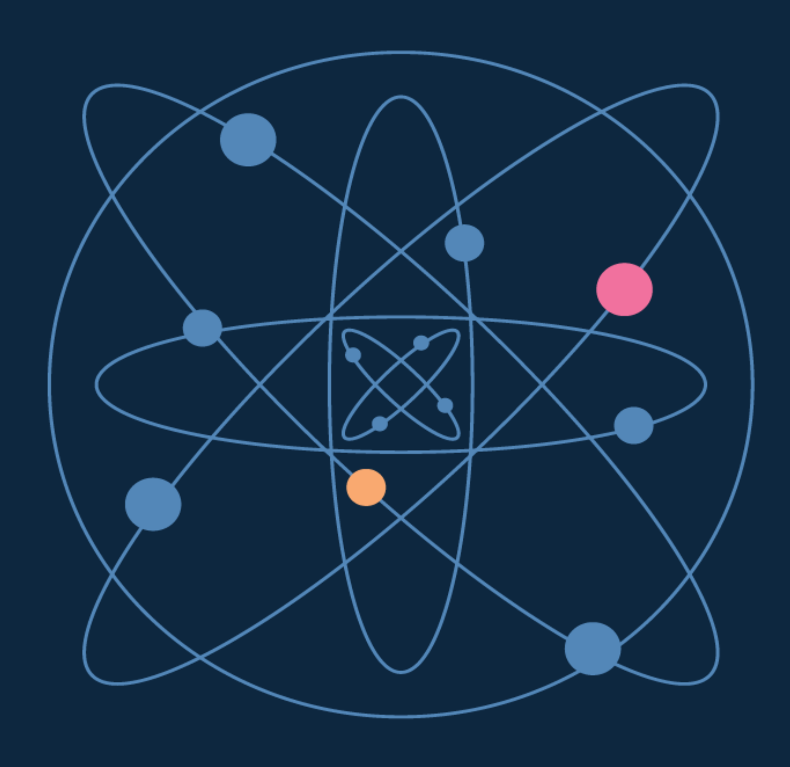
21世纪高等学校计算机系列规划教材



COMPUTER NETWORKS

计算机网络

肖 锋 唐俊勇 主编

清华大学出版社

21 世纪高等学校计算机系列规划教材

计算机网络

肖　锋　唐俊勇　主编

清华大学出版社
北　京

内 容 简 介

计算机网络是计算机与通信技术紧密结合并不断发展的一门学科。本书是一本介绍计算机网络与应用的教材，按照教育部关于计算机及相关专业计算机网络应用的要求，并结合当前计算机网络发展变化而编写。本书既注重计算机网络基础理论的讲解，又注重实践和应用，从先进性和实用性出发，较全面地介绍了计算机网络所涉及的基本理论和应用实践。

本书共分为 11 章，全面系统地介绍了计算机网络体系结构和发展，以及数据通信基础、物理层、数据链路层、局域网技术、网络层、传输层、应用层、网络管理与网络安全、网络系统集成等内容。书中以 OSI 参考模型为基础，突出 TCP/IP 协议栈的常用网络协议，并且包括了虚拟局域网、无线局域网、IPv6 及移动 IP 等新技术和一些最新进展，体现了对新技术的吸收和消化，源源不断地为教学内容补充新鲜血液。

本书可以作为计算机专业、信息技术及电子信息等相关专业本科生、高职生等的网络课程教材，也可以作为相关专业工程技术人员继续教育的培训教材，还可以作为广大网络管理人员或技术人员学习网络知识的参考书。

图书在版编目(CIP)数据

计算机网络/肖锋，唐俊勇主编．—北京：清华大学出版社，2018
(21 世纪高等学校计算机系列规划教材)
ISBN 978-7-302-50044-5

Ⅰ．①计…　Ⅱ．①肖…　②唐…　Ⅲ．①计算机网络－高等学校－教材　Ⅳ．①TP393

中国版本图书馆 CIP 数据核字(2018)第 086641 号

责任编辑：杜　晓
封面设计：傅瑞学
责任校对：李　梅
责任印制：丛怀宇

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编**：100084
社 总 机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
课件下载：http://www.tup.com.cn，010-62770175-4278
印 装 者：三河市铭诚印务有限公司
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　　张**：19　　**字　　数**：458 千字
版　　次：2018 年 5 月第 1 版　　**印　　次**：2018 年 5 月第1 次印刷
印　　数：1～3000
定　　价：49.00 元

产品编号：078273-01

前　言

计算机的产生和发展，彻底改变了人们的工作和生活方式，为人们带来了极大的方便。计算机网络渗透到人类学习与工作的各个方面，为工作和学习提供了许多帮助，担当着越来越重要的角色。Internet的产生与普及、数据资源的共享已经成为21世纪的潮流趋势，学习和了解有关计算机网络及其应用知识十分必要。

进入21世纪，我国高等教育进入前所未有的大发展时期，时代的进步与发展对高等教育质量提出了更高、更新的要求。2001年8月，教育部颁布了《关于加强高等学校本科教学工作，提高教学质量的若干意见》。文件指出，本科教育是高等教育的主体和基础，抓好本科教育是提高整个高等教育质量的重点和关键。随着高等教育的普及和高等学校的扩招，在校大学本科计算机专业学生的人数大幅度上升，对适合21世纪大学本科生学习的计算机相关教材的需求量也将急剧增加，为此，我们组织多名常年讲授计算机网络课程的一线教师，编写了这本适合在校学生和广大计算机爱好者使用的《计算机网络》教材。本书的最大特点是针对学生应用型能力培养的需要，力求理论与实践无缝链接。根据实际需要，介绍有关理论，同时注重应用实践，使学生在掌握基本理论的基础上，具有良好的网络应用和再学习能力。

本书包括11章，各章节讨论主题如下。

第1章：介绍计算机网络的定义、形成、发展、分类与功能、网络的拓扑结构等相关基本概念。

第2章：介绍开放系统互联模型、TCP/IP参考模型、计算机网络体系结构等内容。

第3章：介绍有关数据通信、物理层、多路复用技术以及宽带接入技术。

第4章：介绍数据链路层的基本概念、功能与作用；差错控制的作用和原理、流量控制的原因、原理及方法；面向比特型数据链路层协议——HDLC、PPP等。

第5章：介绍局域网有关知识，包括IEEE 802参考模型体系结构、介质访问控制方法、各种类型以太网、无线局域网以及虚拟局域网等。

第 6 章：介绍网络层相关内容，包括网络层的功能、数据交换方式、IP 协议、IP 路由、无分类编址 CIDR、拥塞控制以及新技术 IPv6 和移动 IP 等。

第 7 章：介绍网络互联主要概念以及网络各层的互联设备，包括中继器、集线器、交换机以及路由器等。

第 8 章：介绍传输层相关内容，包括传输层的功能、TCP 以及 UDP。

第 9 章：介绍 Internet 技术与应用层的相关内容，包括 DNS 域名机制、FTP 服务、E-mail 服务以及 Web 服务等。

第 10 章：介绍网络管理与网络安全的相关内容，包括网络管理的概念、SNMP、网络安全概念以及防火墙、VPN 和网络病毒等。

第 11 章：介绍网络系统集成设计的一般步骤、设计原则和布线规则，通过不同规模企业网络的规划设计案例来说明一般的网络设计过程。

本书可以作为计算机专业、信息技术及电子信息等相关专业本科生、高职生等的网络课程教材，也可以作为相关专业工程技术人员继续教育的培训教材，还可以作为广大网络管理人员或技术人员学习网络知识的参考书。

本书由肖锋和唐俊勇任主编。编写分工为：唐俊勇编写第 1 章和第 10 章；田鹏辉编写第 2 章和第 3 章；曹子建编写第 4 章和第 7 章；肖锋编写第 5 章和第 6 章；王辉编写第 8 章和第 9 章；刁振军编写第 11 章，最后由肖锋负责统稿和定稿。参与本书编写的还有李茹娜、刘宝同、敬亚娇、马启以及冯飞等。

本书在编写与出版过程中，得到了许多老师的关心和帮助，并提出许多宝贵的修改意见，对于他们的关心、帮助和支持，表示十分感谢。

由于计算机网络技术的快速发展，同时编者水平与时间有限，对书中内容的取舍与把握可能不够准确，书中难免存在疏漏与不妥，恳请同行专家和广大读者批评指正。

编　者

2018 年 1 月

目 录

第1章 计算机网络基础

学习目的

帮助学生在了解网络形成与发展历史的基础上，对网络定义、分类和拓扑结构等几个基本问题进行了系统学习，对网络的各种应用以及网络的研究与发展进行了较为系统的讨论，以帮助学生对计算机网络有一个全面和准确的认识，为今后课程的学习打下基础。

学习要求

理解：计算机网络的形成与发展过程。

掌握：计算机网络的定义。

掌握：计算机网络的组成与结构的基本概念。

掌握：计算机网络分类方法及主要类型。

掌握：计算机网络拓扑结构的定义、分类与特点。

了解：计算机网络研究与应用的发展。

1.1 什么是计算机网络

本节首先来简单了解一下网络的产生与发展，然后介绍网络的定义、分类及主要功能。期望对计算机网络有一个初步的认识。

1.1.1 计算机的产生与发展

1. 以单台计算机为中心的联机终端网络

1946年世界第一台电子数字计算机在美国诞生时，计算机技术与通信技术并没有直接联系。20世纪50年代，由于美国军方的需要，麻省理工学院林肯实验室就开始为美国空军设计称为SAGE的半自动化地面防空系统。这实质上是最早计算机与通信技术相结合的尝试。该系统分为17个防区，每个防区的指挥中心装有两台IBM公司的AN/FSQ-7计算机，通过通信线路连接防区内各雷达站、机场、防空导弹和高射炮阵地，形成联机计算机系统，最终实现分布式的防空信息集中处理与控制。SAGE系统最先采用了实现人机交互的显示器，使用了小型计算机形式的前端处理机，制定了1600bps的数据通信规程，并提供了高可靠性的路由选择算法。这个系统最终于1963年建成，被认为是计算机技术和通信技术结合的先驱。

计算机通信技术应用于民用系统方面，最早的是美国航空公司与IBM公司在20世纪50年代初开始联合研究，60年代初投入使用的飞机订票系统SABRE-1。这个系统由一台

中央计算机与整个美国本土内的2000个终端组成。这些终端采用多点线路与中央计算机相连。

在这个阶段,联机终端是一种主要的系统结构形式。所谓终端就是不具有处理和存储能力的计算机。图1.1所示这种以单机互联系统为中心的互联系统,即主机面向终端系统。在这些早期的单台计算机联机网络中,已涉及多种通信技术、多种数据传输设备和数据交换设备。技术上已从单用户系统发展到了远距离的分时多用户系统。虽然联机终端网络在当时的历史条件下已充分显示了计算机与通信相结合的巨大优势,但它仍然有严重的缺点:①主机负荷较重,既要承担通信任务,又要进行数据处理;②通信线路的利用率低,尤其在远距离时,分散的终端都需要独占一条通信线路,不仅通信费用昂贵而且通信线路利用率低;③这种结构属集中控制方式,可靠性低。这期间已经使用了多点通信线路、集中器以及前端处理机。

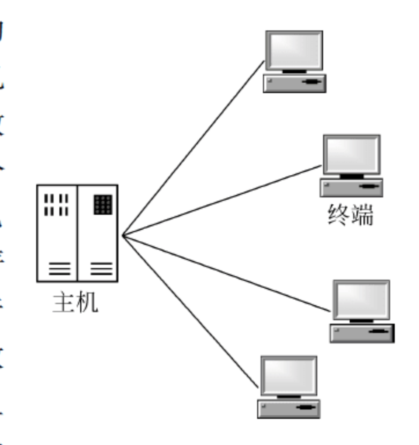


图1.1 面向终端的单机互联系统

2. 多台主计算机通过线路互联的计算机网络

为了克服第一代计算机网络的缺点,提高网络的可靠性和可用性,人们开始研究将多台计算机相互连接的方法。20世纪60年代中期开始,出现、发展了若干台计算机互联的系统,开创了"计算机—计算机"通信的时代。

第二代网络是从20世纪60年代中期到70年代中期,已经形成了将多台主机互联系统相互连接起来,以多处理机为中心的网络,并利用通信线路将多台主机连接起来,为终端用户提供服务,如图1.2所示。

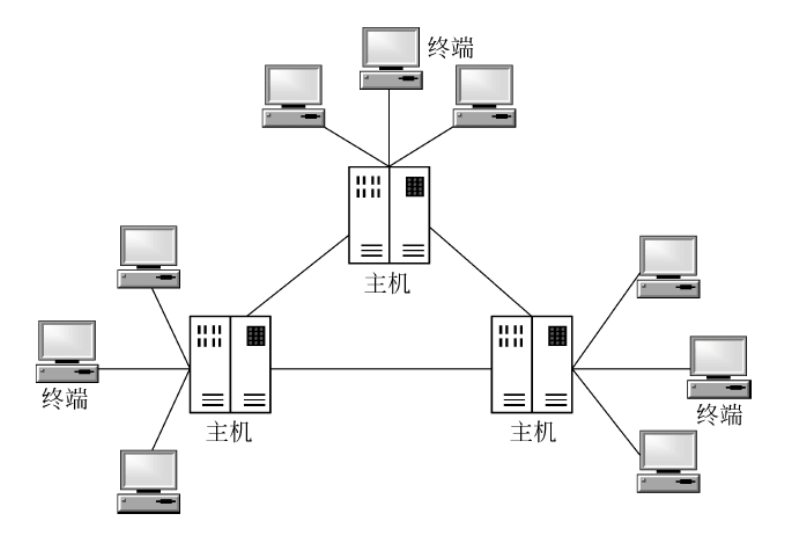


图1.2 多台主机互联系统

这个阶段的典型代表是美国国防部高级研究计划局(Advanced Research Projects Agency)的ARPANET(通常称为ARPA网),标志着现代意义上的计算机网络的出现。1969年由美国国防部高级研究计划局提供经费,联合计算机公司和大学共同研制而发展起来,主要目标是借助于通信系统,使网内各计算机系统间能够相互共享资源,最终导致一个

实验性的4个节点网络开始运行并投入使用。到1973年发展到40个节点,而到了1983年已经达到100个计算机节点,地理上不仅跨越美国本土,而且通过卫星链路连接夏威夷和欧洲的节点。ARPA网所具有的资源共享、分散控制、分组交换、专用的通信控制处理机以及分层的网络协议等特点往往被认为是现代计算机网络的一般特征。所以ARPA网是计算机网络技术发展的一个重要里程碑。

在ARPA网中,将计算机网络分为资源子网和通信子网,如图1.3所示。

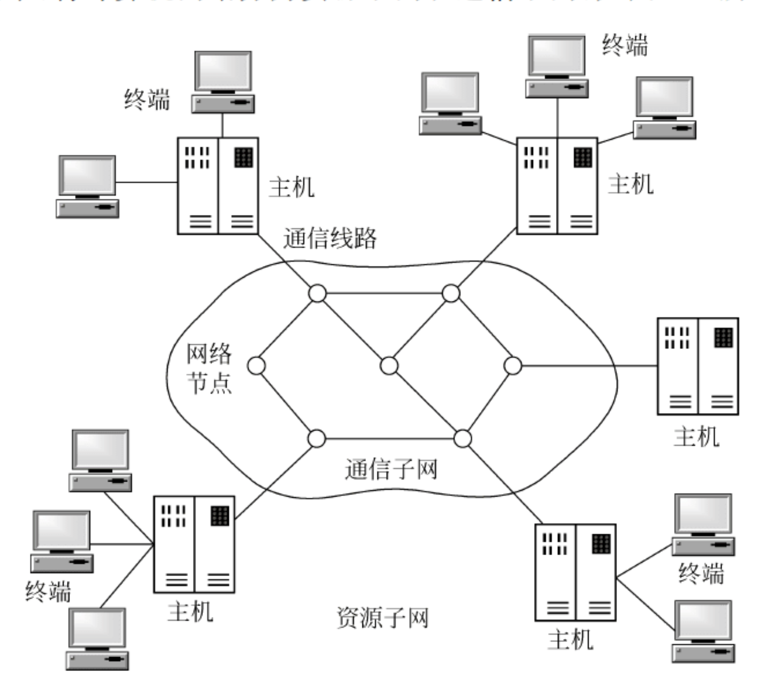


图1.3　资源子网和通信子网

3. 具有统一的网络体系结构、遵循国际标准化协议的计算机网络

在第三代网络出现以前网络是无法实现不同厂家设备互联的,早期,为了霸占市场,各厂家采用自己独特的技术并开发了自己的网络体系结构,当时,IBM发布了SNA(System Network Architecture,系统网络体系结构),DEC公司发布了DNA(Digital Network Architecture,数字网络体系结构)。不同的网络体系结构是无法互联的,所以不同厂家的设备无法达到互联,即使是同一家产品在不同时期也是无法达到互联的,这样就阻碍了大范围网络的发展。后来,为了实现网络大范围的发展和不同厂家设备的互联,1977年国际标准化组织ISO(International Organization for Standardization,ISO)提出一个标准框架——OSI(Open System Interconnection/Reference Model,开放系统互联参考模型),共七层。1984年正式发布了OSI,使厂家设备、协议达到全网互联。

20世纪80年代随着微型计算机的普及与推广应用,计算机局域网开始盛行起来。当时采用的是具有统一的网络体系结构并遵守国际标准的开放式和标准化的网络,它是网络发展的第三阶段。

局域网的发展道路不同于广域网,局域网厂商从一开始就按照标准化、互相兼容的方式展开竞争。现在,局域网厂商大多进入专业化的成熟时期。目前,在一个用户的局域网中,工作站可能是IBM公司的,服务器可能是Compaq公司生产的,网卡可能是3Com公司生产的,集线器可能是DEC公司生产的,而网络上所运行的软件则可能是微软公司的。如果

没有网络标准化,这种高度的兼容性是不可想象的。

4. 因特网时代(Internet)与物联网

进入20世纪90年代后至今都是属于第四代计算机网络。第四代计算机网络是随着数字通信出现和光纤的接入而产生的,其特点:网络化、综合化、高速化及计算机协同能力。

这一阶段因特网(Internet)成为计算机网络领域最引人注目也是发展最快的网络技术。如今Internet已经成为遍布全球的国际性网络,与之相连的网络有数百万个,在网上运行的主机达数亿台,而且它们还在以飞快的速度不断增加。Internet上不仅有分布于世界各地计算机上无穷无尽的信息资源,而且也为网络用户提供各种各样的网络应用服务。1994年开始了商业化应用,利用因特网进行商业活动成为世界经济的一大热点。更高性能的Internet2正在发展之中。可以说Internet的普及与应用是人类由工业社会向信息社会发展的重要标志。

Internet的网络体系结构采用TCP/IP协议集。TCP/IP协议集由一组以TCP和IP为代表的协议构成,采用四层结构。它简单实用,能满足不同服务需求的数据传输,已成为业内公认的标准。使用TCP/IP可方便地将不同类型的主机和网络互联,原则上任何计算机只要遵守TCP/IP协议,都能按一定的规则接入Internet。现在,快速网络接入Internet的方式也在不断地诞生,如ISDN、ADSL、DDN、FDDI和ATM网络等。

目前,基于光纤通信技术和宽带城域网与无线网技术,以及移动网格计算、网络多媒体计算、网络并行计算、网格计算与存储区域网络等正成为网络应用与研究的热点问题。

1.1.2 计算机网络的定义

在给出计算机网络的定义之前,先来回顾一下我们平常所说的"网络"概念。"网络"通常是指为了达到某种目标而以某种方式联系或组合在一起的对象或物体的集合。如我们日常生活中四通八达的交通系统、供水或供电系统、邮政系统等都是某种形式的网络。

在计算机网络发展过程中,人们在不同阶段或从不同角度对计算机网络提出了不同的定义,其中比较典型的有3类:广义的观点、资源共享的观点和用户透明性的观点。

1. 广义的观点

计算机网络是以实现远程通信为目的,一些互联的、独立自治的计算机的集合。这种观点提出得比较早,它以计算机相互间的数据传输为主要目的,资源共享能力弱,对计算机网络的认识主要停留在计算机通信网络的层面,是计算机网络的初级阶段,是一种比较早期的观点。

2. 资源共享的观点

计算机网络是把地理位置上分散的,各自具有独立功能的计算机系统,以能够相互共享资源的方式连接起来的集合。计算机网络具有如下特征。

(1) 计算机通信的目的是为共享硬件、软件以及信息资源。

(2) 各计算机功能独立,地域可以分散。

(3) 计算机网络应具有网络操作系统,遵循统一的网络协议。

3. 用户透明性的观点

计算机网络是一组相互连接在一起的计算机系统的集合，使得整个网络像一个大的计算机系统一样，因此它对用户是透明的。这种观点主要关注网络作为一种分布式系统，从而以内部资源分布与资源调度等技术实现应该对用户透明的角度来描述计算机网络，使整个网络像一个大的计算机系统一样对用户是透明的。

然而，从目前计算机网络的发展现状与特征来看，我们认为从资源共享的角度理解计算机网络比较准确和全面。

从资源共享的角度来看，计算机网络是指将地理位置不同且功能相对独立的多个计算机系统通过通信线路相互连在一起、遵循共同的网络协议、由专门的网络操作系统进行管理，以实现资源共享的系统。

这主要包含了三层含义：建立计算机网络的目的就是实现计算机资源共享；彼此独立则强调在网络中，计算机之间不存在明显的主从关系，即网络中的计算机不具备控制其他计算机的能力，每台计算机都具有独立的操作系统；联网计算机之间的通信必须遵循共同的网络协议。

“地理位置不同”是指计算机网络中的计算机通常都处于不同的地理位置。例如，当用户通过互联网访问某种网络服务时，被访问的主机在地理上往往是不可见的，主机可能位于不同的城市、省份乃至不同的国家，所以这些被访问的主机有时被称为远程主机。事实上，在绝大部分情况下大家甚至不知道也不需要知道这个被访问机器所处的确切位置。

“功能相对独立”是指相互连接的计算机之间不存在互为依赖的关系。作为各自独立的计算机系统，它们具有各自独立的软件和硬件。任何一台计算机既可以联网工作，也可以脱离网络和网络中的其他计算机独立工作。

1.1.3 计算机网络的功能

计算机网络技术使计算机的作用范围和其自身的功能有了突破性的发展。计算机网络虽然各种各样，但作为计算机网络都应具有如下功能。一般来说，计算机网络具有的功能，又称为服务。其中最主要的功能是资源共享和数据通信。

1. 资源共享

资源共享是网络的基本功能之一。资源共享包括硬件资源的共享，如大型主机、大容量磁盘、光盘库、打印机、网络通信设备和通信线路与服务器硬件等；也包括软件资源的共享，如网络操作系统、数据库管理系统、网络管理系统、应用软件、开发工具和服务器软件等。资源共享功能不仅使得网络用户可以克服地理位置上的差异，共享网络中的资源，充分提高资源的利用率，还可以避免重复投资和劳动，使系统的整体性价比得到改善。资源共享是计算机网络产生的主要原动力。通过资源共享，可使网络中各处的资源互通有无、分工协作，从而大大提高系统资源的利用率。

2. 数据通信

通信即在计算机之间传送信息，是计算机网络最基本的功能。通过计算机网络使不同地区的用户可以快速和准确地相互传送信息，这些信息包括数据、文本、图形、动画、声音和视频等。用户还可以收发 E-mail、播放 VOD(视频点播)和拨打 IP 电话等。

3. 进行数据信息的集中和综合处理

将分散在各地计算机中的数据资料适时集中或分级管理，并经综合处理后形成各种报表，提供给管理者或决策者分析和参考，如自动订票系统、政府部门的计划统计系统、银行财政及各种金融系统、数据的收集和处理系统、地震资料收集与处理系统、地质资料采集与处理系统等。今天，我们之所以能够足不出户，通过网络来实现飞机票、火车票的预订，在出差到外地之前就能完成饭店、住宿的预订，正是得益于网络所提供的信息集中和综合处理功能。

4. 均衡负载，相互协作

当某一个计算中心的任务很重时，可通过网络将此任务传递给空闲的计算机去处理，通过协同操作和并行处理等方式来分担负载。不仅可以使网内各计算机负载均衡，还可以充分利用计算机网络内的空闲资源来提高整个系统的处理能力。对于一个用户访问量非常大的热点网站，当它的单台服务器不能满足用户的访问需求时，可以考虑以多台服务器所构成的集群(Cluster)来实施负载均衡，为用户提供更加有效的服务。在一些幅员辽阔的国家，还可以利用时差来均衡不同网络节点上的日夜负荷。

5. 提高了系统的可靠性和可用性

当网中的某一处理机发生故障时，可由别的路径传输信息或转到别的系统中代为处理，以保证用户的正常操作，不因局部故障而导致系统的瘫痪。又如，某一数据库中的数据因处理机发生故障而消失或遭到破坏时，可从另一台计算机的备份数据库中调来进行处理，并恢复遭破坏的数据库，从而提高系统的可靠性和可用性。

6. 进行分布式处理

对于综合性的大型问题可采用合适的算法，将任务分散到网中不同的计算机上进行分布式处理。特别是对当前流行的局域网更有意义，利用网络技术将计算机连成高性能的分布式计算机系统，使它具有解决复杂问题的能力。例如，在一个分布式的气象信息处理系统中，可以调用遍布在十分辽阔地域范围内的各计算机协同工作，对所获得的卫星气象数据进行快速、及时处理，以得到准确的气象信息。

1.1.4 计算机网络的应用与发展

随着现代信息社会进程的推进以及通信和计算机技术的迅猛发展，计算机网络的应用日益多元化，打破了空间和时间的限制，几乎深入社会的各个领域。归纳起来有三个明显的特征：①应用的多样性；②新应用的产生速度加快；③应用的领域日趋广泛。下面列举一些典型的网络应用。

1. 方便的信息检索

计算机网络使我们的信息检索变得更加高效、快捷，通过网上搜索、WWW 浏览、FTP 下载，可以非常方便地从网络上获得所需的信息和资料。网上图书馆更以其信息容量大、检索方便的优势赢得了人们的青睐。

2. 现代化的通信方式

网络上使用最为广泛的电子邮件目前已经成为一种最为快捷、廉价的通信手段。人们在几分，甚至几秒内就可以把信息发送给对方，信息的表达形式不仅可以是文本，还可以

是声音和图片。其低廉的通信费用更是其他通信方式（如信件、电话、传真等）所不能相比的。同时，利用网络可以实现IP电话，将语音和数据网络进行集成，利用IP作为传输协议，通过网络将语音集成到IP网络上来，在基于IP的网络上进行语音通信，节省长途电话费用。

3. 办公自动化

通过将一个企业或机关的办公计算机及其外部设备连成网络，既可以节约购买多个外部设备的成本，又可以共享许多办公数据，并且可对信息进行计算机综合处理与统计，避免了许多单调重复性的劳动。

4. 电子商务与电子政务

计算机网络还推动了电子商务与电子政务的发展。企业与企业之间、企业与个人之间可以通过网络来实现贸易、购物；政府部门则可以通过电子政务工程实施政务公开化，审批程序标准化，提高了政府的办事效率并使之更好地为企业或个人服务。

5. 企业的信息化

通过在企业中实施基于网络的管理信息系统（MIS）和资源制造计划（ERP），可以实现企业的生产、销售、管理和服务的全面信息化，从而有效地提高生产率。医院管理信息系统、民航及铁路的购票系统、学校的学生管理信息系统等都是管理信息系统的实例。

6. 远程教育与E-Learning

网络提供了新的实现自我教育和终身教育的渠道。基于网络的远程教育、网络学习使得我们可以突破时间、空间和身份的限制，方便地获取网络上的教育资源并接受教育。

7. 丰富的娱乐和消遣

网络不仅改变了我们的工作与学习方式，也带来了新的丰富多彩的娱乐和消遣方式，如网上聊天、网络游戏、网上电影院、视频点播等。

8. 军事指挥自动化

基于C^4I的网络应用系统，把军事情报采集、目标定位、武器控制、战地通信和指挥员决策等环节在计算机网络基础上联系起来，形成各种高速高效的指挥自动化系统，是现代战争和军队现代化不可缺少的技术支柱，这种系统在公安武警、交警、火警等指挥调度系统中也有广泛应用。

1.2　计算机网络的组成与结构

计算机网络的组成可以从物理组成和逻辑组成两个角度进行理解。

1.2.1　计算机网络的物理组成

计算机网络的物理组成主要是从资源构成的角度讲，由网络硬件系统和网络软件系统组成。

1. 硬件系统

局域网的硬件系统是指构成局域网的所有物理设备总和。主要包括计算机设备和通信

传输设备，这些设备按照功能和在网络中的作用可以分为服务器、客户机、网络适配器（网卡）、通信设备、连接设备和连接介质等，主要包括以下部分。

（1）服务器：一般为高性能计算机，用于网络管理，运行应用程序，处理各网络工作站成员的信息请求等，并可连接一些外部设备，如打印机、CD-ROM、调制解调器等。根据其作用的不同分为文件服务器、应用程序服务器和数据库服务器等。

（2）工作站：由服务器进行管理和提供服务的、连入网络的任何计算机都属于工作站，其性能一般低于服务器。个人计算机接入因特网后，在获取因特网服务的同时，其本身就成为一台因特网上的工作站。

（3）传输介质：用于网络设备之间通信连接的网络电缆。常用的传输介质有双绞线、细同轴电缆、粗同轴电缆、光缆等。

（4）网卡：也称网络适配器，在局域网中用于将计算机与网络相连接的设备。

（5）交换机（Switch）：用于处理与高带宽相关协议配套的多层设备，具有数据转发、过滤等功能，适用于数据量大、通信频繁的网络。

（6）集线器（Hub）：用来连接多台计算机或局域网络的设备。对所连接的网络介质上的信号有再生和放大的作用，可使所连接的介质长度达到最大有效长度，需要有电源才能工作。

（7）调制解调器：用来实现模拟和数字信号之间转换的设备，由于电信网采用模拟信号传输，而计算机内部信号为数字信号，所以该设备就成为计算机拨号上网的必备设备。

（8）路由器（Router）：一种网络层设备，可互联局域网和广域网，并且当网络上两端点间存在多条通路时，路由器可以提供交通控制和筛选功能，选择信息通路。

（9）光纤：玻璃制成的传输介质，通常用于高速通信线路的连接。

（10）终端（Terminal）：没有中央处理器（CPU）的网络工作站，在早期网络中采用。现代网络中，可以独立工作的终端称为“客户机”。

2. 软件系统

计算机网络的软件系统包括网络通信协议、网络操作系统和网络应用软件。

1）网络通信协议

计算机与计算机之间的通信离不开通信协议，通信协议实际上是一组规定和约定的准则。两台计算机在通信时必须约定好本次通信做什么，是进行文件传输，还是发送电子邮件，怎样通信，什么时间通信等。因此，通信双方要遵从相互可以接受的协议（相同或兼容的协议）才能进行通信。如目前因特网上使用的 TCP/IP 协议等，任何计算机联入网络后只要运行 TCP/IP 协议，就可以访问因特网。

2）网络操作系统

网络操作系统（Network Operation System，NOS）是指能使网络上多台计算机方便而有效地共享网络资源，为用户提供所需各种服务的操作系统软件。网络操作系统除了具备单机操作系统所需的功能外，还应提供高效可靠的网络通信能力和提供多项网络服务功能，如远程管理、文件传输、电子邮件、远程打印等。

目前，最为流行的网络操作系统为 Windows NT、UNIX、Linux 等。

3）网络应用软件

网络应用软件是构建在网络操作系统上的应用程序，不同的应用软件，可以满足网络用

户不同的需求,例如,网络数据库软件、网络通信软件等。

1.2.2 计算机网络的逻辑组成

但是,从计算机网络的设计与实现角度看,更多的是从功能角度去看待计算机网络的组成,并从功能上将计算机网络逻辑划分为资源子网和通信子网,如图 1.4 所示。

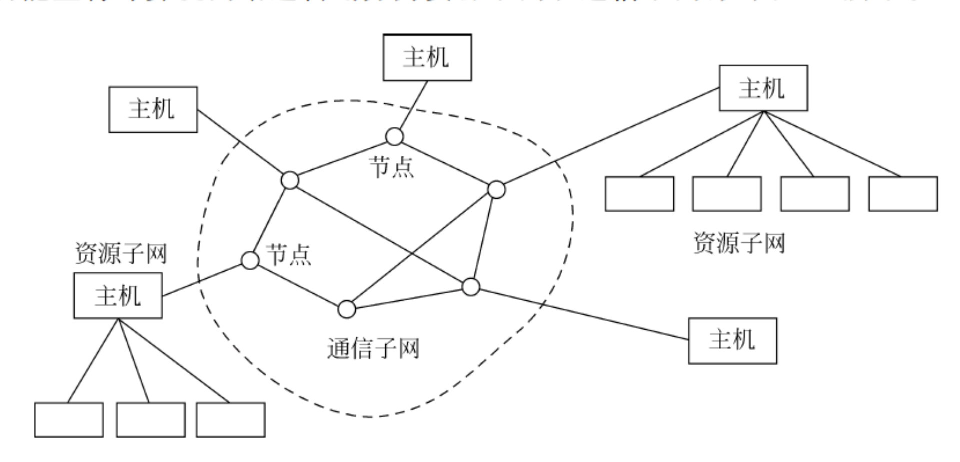


图 1.4 资源子网及通信子网的划分

1. 资源子网

资源子网负责全网的数据处理业务,并向网络用户提供各种网络资源和网络服务。资源子网主要由主机、终端以及相应的 I/O 设备、各种软件资源和数据资源构成。

主机(Host)可以是大型机、中型机、小型机、工作站或微型机,它通过高速通信线路与通信控制处理机相连。主机系统拥有各种终端用户要访问的资源,它负担着数据处理的任务。

终端(Terminal)是用户进行网络操作时所使用的末端设备,它是用户访问网络的界面。终端设备的种类很多,如电传打字机、CRT 监视器加键盘,另外还有网络打印机、传真机等。终端设备可以直接或者通过通信控制处理机和主机相连。

2. 通信子网

通信子网是由负责数据通信处理的通信控制处理机和传输链路组成的独立的数据通信系统,它负担着全网的数据传输、转接、加工和变换等通信处理工作,它主要包括通信线路、网络连接设备、网络通信协议、通信控制软件等。

随着计算机网络技术的发展特别是微型计算机和路由设备的广泛使用,现代网络中的通信子网与资源子网内部已经发生了显著的变化。在资源子网中,大量的微型计算机通过局域网(包括校园网、企业网或 ISP 提供的接入网等)连入广域网;在通信子网中,用于实现广域网与广域网之间互联的通信控制处理机普遍采用了被称为核心路由器的路由设备;在资源子网和通信子网的边界,局域网与广域网之间的互联也采用了路由设备,并将这些路由设备称为接入路由器或边界路由器。图 1.5 给出了现代计算机网络结构的简单示意图。

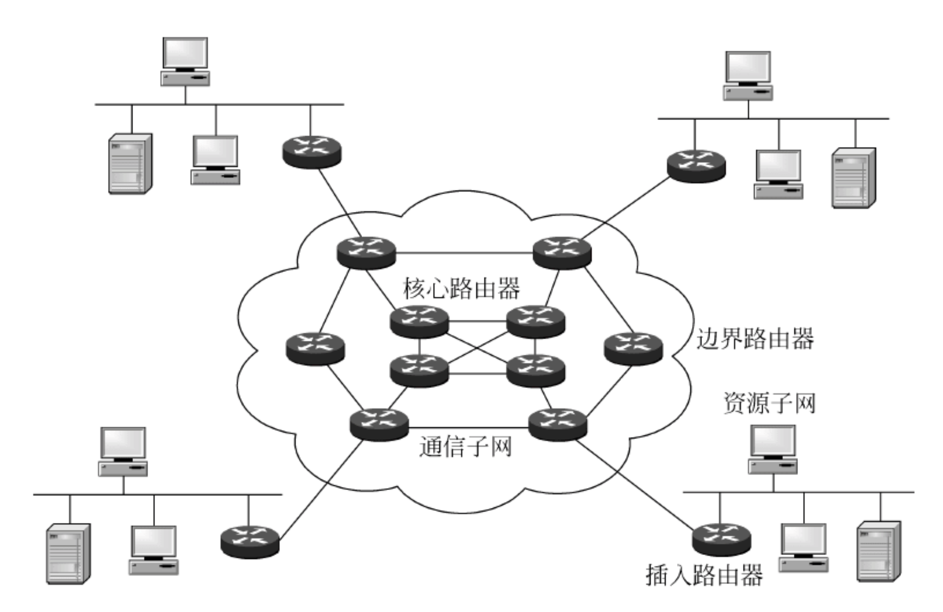


图 1.5　现代计算机网络结构的简单示意图

1.3　计算机网络的分类

计算机网络的分类方法很多，按照不同分类标准，可以将计算机网络分为多种不同类型。

1. 按通信传输介质分类

按通信传输介质不同可分为有线网络和无线网络。所谓有线网络，是指采用有形的传输介质，如双绞线、同轴电缆、光纤等组建的网络，而使用微波、红外线等无线传输介质作为通信线路的网络就属于无线网络和卫星网络等。

2. 按使用网络的对象分类

按使用网络的对象不同可分为专用网和公用网。专用网一般由某个单位或部门组建，使用权限属于单位或部门内部所有，不允许外单位或部门使用，如银行系统的网络。而公用网由电信部门组建，网络内的传输和交换设备可提供给任何部门与单位使用，如 Internet。

3. 按网络的拓扑结构分类

网络的拓扑结构是指一个网络的通信链路和节点的几何排列或物理布局图形分支，它是从图论演变过来的。在网络中常用的拓扑结构有星形拓扑、环形拓扑、总线型拓扑。网形拓扑、树形拓扑等都是上述 3 类拓扑结构的扩展。详细内容见后续章节。

4. 按数据传输方式分类

按数据传输方式的不同，计算机网络又可以分为“广播网络”和“点对点网络”两大类。

广播网络(Broadcasting Network)中的计算机或设备都共享一个公共通信信道。当一台计算机利用共享信道发送报文分组时，所有其他计算机都会“收听到”这个分组。接收到该分组的计算机将通过检查目的地址来决定是否接收该分组。局域网和城域网基本上都是采用广播式通信信道与技术，而广域网中的无线网络、卫星网络也采用广播式通信信道与技术。

广播网络中的传输方式目前有以下3种方式。

(1) 单播(Unicast):发送的信息中包含明确的目的地址,所有节点都检查该地址。如果与自己的地址相同,则处理该信息;如果不同,则忽略。

(2) 组播(Multicast):将信息传输给网络中的部分节点。

(3) 广播(Broadcast):在发送的信息中使用一个指定的代码标识目的地址,将信息发送给所有的目标节点。当使用这个指定代码传输信息时,所有节点都接收并处理该信息。

点到点传播是指网络中每两台主机、两台节点交换机之间或主机与节点交换机之间都存在一条物理信道,即每条物理线路连接一对计算机。机器(包括主机和节点交换机)沿某信道发送的数据确定无疑地只有信道另一端的唯一一台机器收到。与广播网络相比,点对点网络具有以下主要特点。

(1) 每条物理线路连接一对计算机。假如两台计算机之间没有直接连接的线路,那么它们之间的分组传输就要通过中间节点的接收、存储与转发,直至目的节点。

(2) 由于连接多台计算机之间的线路结构可能是复杂的,因此从源节点到目的节点可能存在多条路由。决定分组从通信子网的源节点到目的节点的路由需要有路由算法。

(3) 采用分组存储转发与路由选择机制是点对点网络与广播网络的重要区别之一。广域网基本上都属于点对点网络。

5. 按网络的覆盖范围分类

计算机网络技术按照其覆盖的地理范围进行分类,可以很好地反映不同类型网络的技术特征。由于网络覆盖的地理范围不同,它们所采用的技术也就不同,因而形成了不同的网络技术特点与网络服务功能。按覆盖地理范围的大小,可以把计算机网络分为局域网、城域网和广域网,如表1.1所示。

表1.1　计算机网络的一般分类

网络分类	分布距离	跨越地理范围	带　宽
局域网	10m	房间	10Mbps～xGbps
	200m	建筑物	
	2km	校园内	
城域网	100km	城市	64Kbps～xGbps
广域网	1000km	国家、洲或洲际	64Kbps～625Mbps

1) 局域网

局域网(Local Area Network,LAN)分布于一个间房、每个楼层、整栋楼及楼群之间等,范围一般在2km以内,最大距离不超过10km,如图1.6所示。它是在小型计算机和微型计算机大量推广使用之后逐渐发展起来的。一方面,它容易管理与配置;另一方面,它容易构成简洁整齐的拓扑结构。局域网速率高,延迟小,传输速率通常为10Mbps～2Gbps。因此,网络节点往往能对等地参与对整个网络的使用与监控。再加上成本低、应用广、组网方便及使用灵活等特点,深受用户欢迎,是目前计算机网络技术发展中最活跃的一个分支。局域网的物理网络通常只包含物理层和数据链路层。

局域网主要用来构建一个单位的内部网络,例如,办公室网络、办公大楼内的局域网、学校的校园网、工厂的企业网、大公司及科研机构的园区网等。局域网通常属于单位所有,单

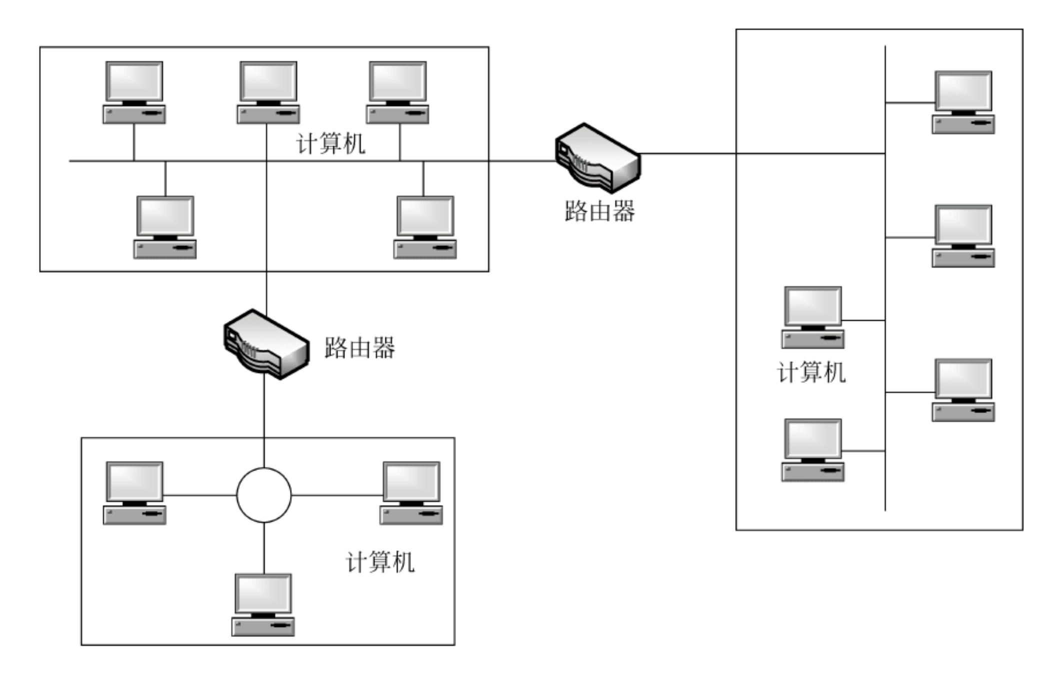


图 1.6 局域网

位拥有自主管理权，以共享网络资源和协同式网络应用为主要目的。

局域网的主要特点：

(1) 适应网络范围小；

(2) 传输速率高；

(3) 组建方便，使用灵活；

(4) 网络组建成本低；

(5) 数据传输错误率低。

局域网按照采用的技术、应用范围和协议标准的不同，可以分为共享局域网和交换局域网。局域网发展迅速，应用日益广泛，是目前计算机网络中最活跃的分支。

2) 城域网

城域网(Metropolitan Area Network，MAN)是介于广域网与局域网之间的一种大范围的高速网络，它的覆盖范围通常为几千米至几十千米，传输速率为 2Mbps～xGbps，如图 1.7 所示。

随着使用局域网带来的好处，人们逐渐要求扩大局域网的范围，或者要求将已经使用的局域网互相连接起来，使其成为一个规模较大的城市范围内的网络。因此，城域网设计的目标是要满足几十千米范围内的大量企业、机关、公司与社会服务部门的计算机联网需求，实现大量用户、多种信息传输的综合信息网络。城域网主要是指在大中型企业集团、ISP、电信部门、有线电视台和政府构建的专用网络与公用网络。

城域网的主要特点：

(1) 适合比 LAN 大的区域(通常用于分布在一个城市的大校园或企业之间)；

(2) 比 LAN 速度慢，但比 WAN 速度快；

(3) 昂贵的设备；

(4) 中等错误率。

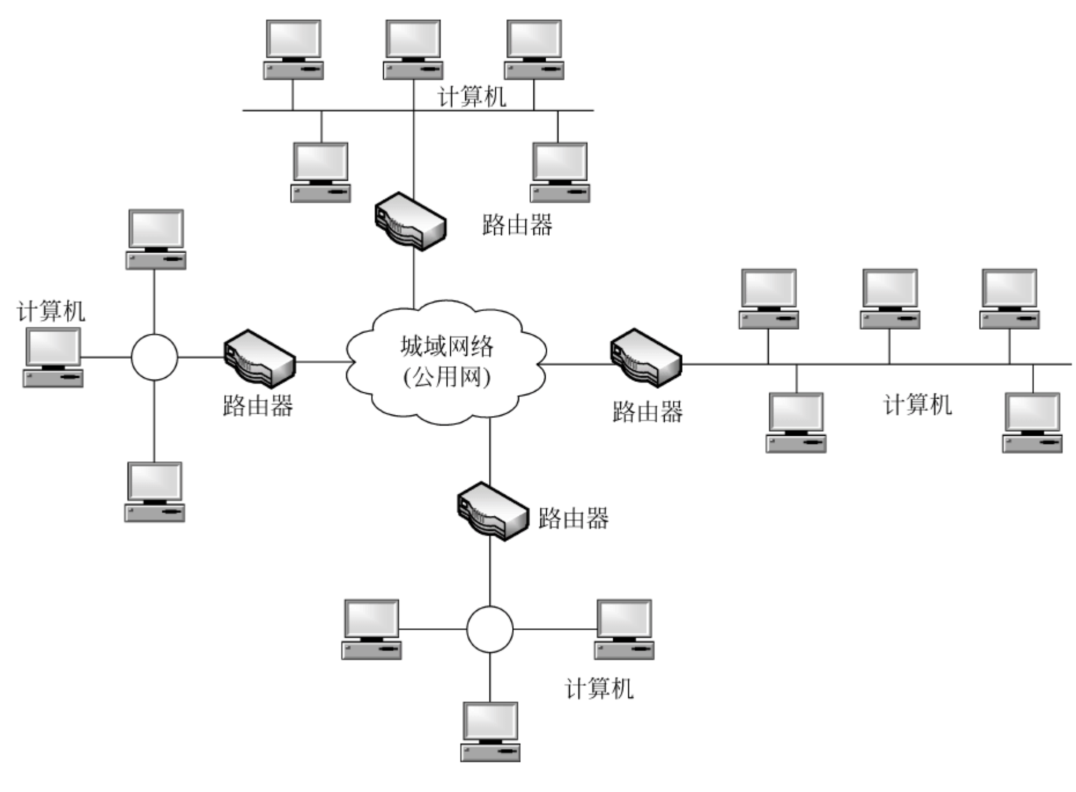


图 1.7　城域网

3）广域网

广域网(Wide Area Network,WAN)的覆盖范围很大,几个城市、一个国家、几个国家甚至全球都属于广域网的范畴,从几十千米到几千千米或几万千米。

此类网络起初是出于军事、国防和科学研究的需要。例如,美国国防 ARPANET 网络,1971 年在全美推广使用并已延伸到世界各地。由于广域网分布距离远,其速率要比局域网低得多。另外在广域网中,网络之间连接用的通信线路大多租用专线,当然也有专门铺设的线路。物理网络本身往往包含了一组复杂的分组交换设备,通过通信线路连接起来,构成网状结构。由于广域网一般采用点对点的通信技术,所以必须解决寻径问题,这也是广域网的物理网络中心包含网络层的原因。

互联网在范畴上属于广域网。但它并不是一种具体的物理网络技术,它是将不同的物理网络技术按某种协议统一起来的一种高层技术,是广域网与广域网、广域网与局域网、局域网与局域网之间的互联,形成了局部处理与远程处理、有限地域范围资源共享与广大地域范围资源共享相结合的互联网。目前,世界上发展最快、最热门的互联网就是 Internet,它是世界上最大的互联网。国内这方面的代表主要有中国电信的 CHINANET 网、中国教育科研网(CERNET)、中国科学院系统的 CSTNET 和金桥网(GBNET)等。

广域网的主要特点：

(1) 规模可以与世界一样大小；

(2) 一般比 LAN 和 MAN 慢很多；

(3) 网络传输错误率最高；

(4) 昂贵的网络设备。

1.4 计算机网络的拓扑结构

拓扑是从数学图论演变而来的，是拓扑学中一种研究与大小、形状无关的点、线、面关系的方法。在计算机网络中也引入了网络拓扑的概念，即忽略具体设备，把工作站、服务器、集线器和路由器等网络单元抽象为点，也称网络节点，把网络中的电缆、双绞线、光纤等通信介质抽象为线。这样从拓扑学的观点看计算机和网络系统，就形成了点和线所组成的几何图形，抽象出网络系统的具体结构。这种采用拓扑学方法抽象出的网络结构称为计算机网络的拓扑结构，它反映了网络中各实体之间的结构关系。

网络拓扑结构图是理解和研究网络的结构与分布的语言。网络拓扑结构反映了网络连接关系的本质，而且还排除了一些没有反映网络本质特征的细节，例如，网络连接所使用的缆线类型和网络主机所使用的操作系统等。网络拓扑结构对整个网络的设计、功能、可靠性、费用及维护等方面有着重要的影响。从某种意义上说，网络拓扑结构图就是网络建设的蓝图。

1.4.1 计算机网络拓扑的基本构型

网络拓扑是指从计算机网络中抽象出来的网络节点与线构成的网络几何形状，或者说它在物理上的连通性。网络的拓扑结构按几何形状主要有以下几种基本构型：星形拓扑、总线型拓扑、环形拓扑、树形拓扑、网状拓扑。所谓网络节点，就是指在网络中独立进行工作的设备。网络节点可能是诸如服务器、工作站等网络主机，也可能是诸如路由器、交换机、集线器、网卡等网络连接设备，如图 1.8 所示。

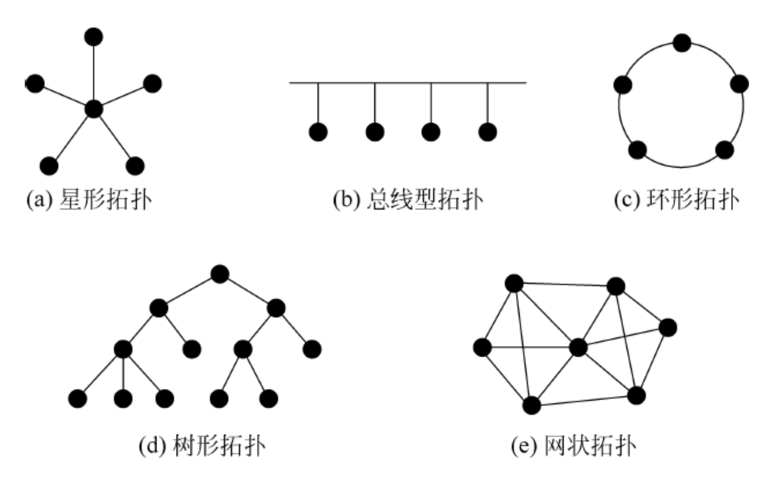


图 1.8 网络拓扑的基本构型

1.4.2 计算机网络拓扑的分类

根据通信子网中通信信道类型，网络拓扑可以分为两类：广播信道通信子网的拓扑与点对点线路通信子网的拓扑。

在采用广播信道通信子网中,一个公共的通信信道被多个网络节点共享。采用广播信道通信子网的基本拓扑结构有 4 种：总线型、环形、树形、无线通信与卫星通信型。这一类主要以局域网的总线型拓扑结构为代表。

在采用点对点线路通信子网中,每条物理线路连接一对节点。采用点对点线路通信子网的基本拓扑结构也有 4 种：网状、星形、环形与树形。目前实际存在与使用的广域网结构,大多采用网状拓扑结构。

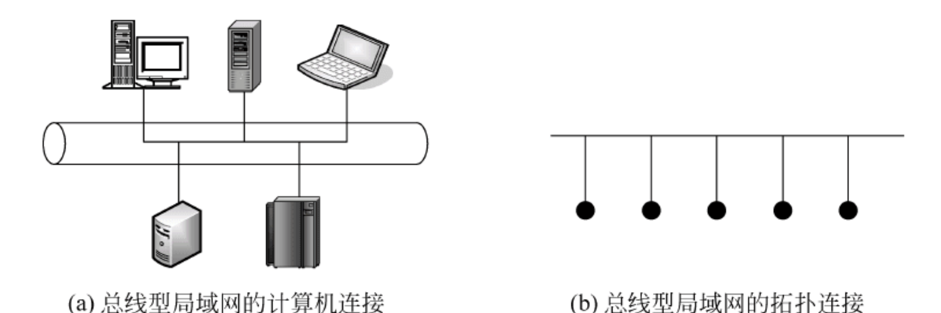
1.4.3 总线型拓扑结构

总线型拓扑结构是局域网主要的拓扑结构之一。它采用单根数据传输线作为通信介质,所有的节点都通过相应的硬件接口(如网卡)直接连接到通信介质,而且能被所有其他的节点接受。图 1.9 所示为总线型拓扑结构。

总线型拓扑结构中的节点为服务器或工作站,通信介质为同轴电缆或双绞线。

由于所有的节点共享一条公用的传输链路,所以一次只能由一个设备传输。这样就需要某种形式的访问控制策略,来决定下一次哪一个节点可以发送。一般情况下,总线型网络采用载波监听多路访问/冲突检测(CSMA/CD)控制策略。

总线型网络信息发送的过程为：发送时,发送节点对报文进行分组,然后一次一个地址依次发送这些分组,有时要与其他工作站传来的分组交替地在通信介质上传输。当分组经过各节点时,目标节点将识别分组的地址,然后将属于自己的分组内容复制下来。



(a) 总线型局域网的计算机连接　　(b) 总线型局域网的拓扑连接

图 1.9　总线型拓扑结构

总线型拓扑结构在局域网中得到了广泛应用,主要优点如下。

(1) 布线容易,电缆用量小。总线型网络中的节点都连接在一个公共的通信介质上,所以需要的电缆长度短,减少了安装费用,易于布线和维护。

(2) 可靠性高。总线结构简单,从硬件观点来看,十分可靠。

(3) 易于扩充。在总线型网络中,如果要增加长度,可通过中继器加上一个附加段；如果需要增加新节点,只需在总线的任何点将其接入。

(4) 易于安装。总线型网络的安装比较简单,对技术要求不是很高。

总线型拓扑结构虽然有许多优点,但也有自己的如下局限性。

(1) 故障诊断困难。虽然总线拓扑简单,可靠性高,但故障检测却不容易。因为具有总线型拓扑结构的网络不是集中控制,故障检测需要在网上各个节点进行。

(2) 故障隔离困难。对于介质的故障,不能简单地撤销某个工作站,这样会切断整段

网络。

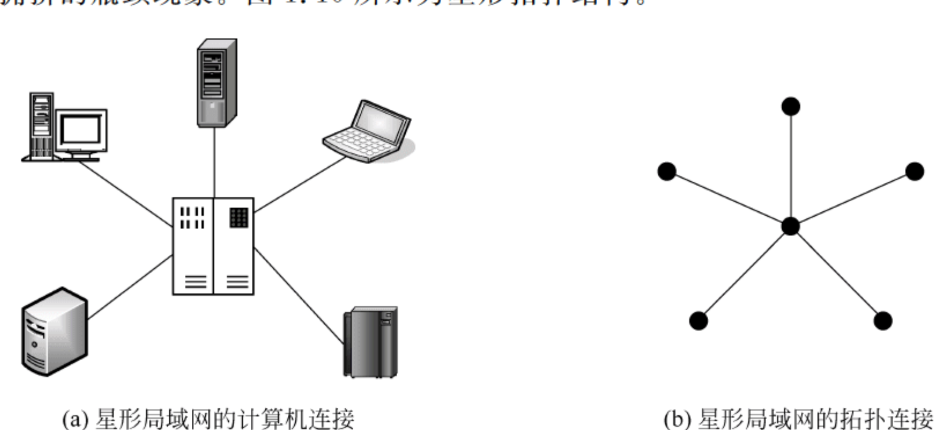
(3) 中继器配置较麻烦。在总线的干线基础上扩充时,可利用中继器,需要重新设置,包括电缆长度的裁剪、终端匹配器的调整等。

(4) 通信介质或中间某一接口点出现故障,整个网络随即瘫痪。

(5) 终端必须是智能的。因为接在总线上的节点有介质访问控制功能,因此必须具有智能,从而增加了站点的硬件和软件费用。

1.4.4 星形拓扑结构

星形拓扑结构由中央节点和通过点到点链路连接到中央节点的各节点组成。利用星形拓扑结构的交换方式有电路交换和报文交换,尤以电路交换更为普遍。一旦建立了通道连接,可以没有延迟地在连通的两个节点之间传送数据。工作站到中央节点的线路是专用的,不会出现拥挤的瓶颈现象。图 1.10 所示为星形拓扑结构。



(a) 星形局域网的计算机连接　　(b) 星形局域网的拓扑连接

图 1.10　星形拓扑结构

星形拓扑结构中,中央节点为集线器(Hub)或局域网交换机,其他外围节点为服务器或工作站;通信介质为双绞线或光纤。

星形拓扑结构被广泛地应用于网络中集中于中央节点的场合。由于所有节点往外传输都必须经过中央节点来处理,因此,对中央节点的要求比较高。

星形拓扑结构信息发送的过程为:某一工作站有信息发送时,将向中央节点申请,中央节点响应该工作站,并将该工作站与目的工作站或服务器建立会话。此时,就可以进行无延时的会话了。

在交换式局域网(Switched LAN)出现后,才真正出现了物理结构与逻辑结构统一的星形拓扑结构。交换式局域网的中心节点是局域网交换机。在典型的交换式局域网中,节点可以通过点对点线路与局域网交换机连接。局域网交换机可以在多对节点之间建立并发连接。

星形拓扑结构的优点如下。

(1) 可靠性高。在星形拓扑结构中,每个连接只与一个设备相连,因此,单个连接的故障只影响一个设备,不会影响全网。

(2) 方便服务。中央节点和中间接线都有一批集中点,可方便地提供服务和进行网络

重新配置。

(3) 故障诊断容易。如果网络中的节点或者通信介质出现问题,只会影响到该节点或者通信介质相连的节点,不会涉及整个网络,从而比较容易判断故障的位置。

星形拓扑结构虽有许多优点,但也有如下缺点。

(1) 扩展困难,安装费用高。增加网络新节点时,无论有多远,都需要与中央节点直接连接,布线困难且费用高。

(2) 对中央节点的依赖性强。星形拓扑结构网络中的外围节点对中央节点的依赖性强,如果中央节点出现故障,则全部网络不能正常工作。

1.4.5 环形拓扑结构

环形拓扑结构是一个像环一样的闭合链路,在链路上有许多中继器和通过中继器连接到链路上的节点。也就是说,环形拓扑结构网络是由一些中继器和连接到中继器的点到点链路组成的一个闭合环。在环形网中,所有的通信共享一条物理通道,即连接网中所有节点的点到点链路。图 1.11 所示为环形拓扑结构。

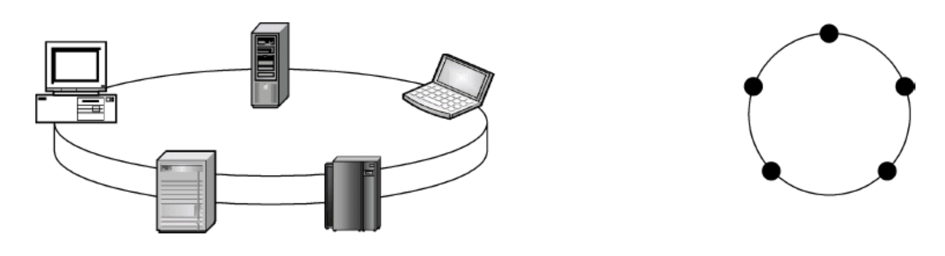


图 1.11　环形拓扑结构

其中,每台中继器通过单向传输链路连接到另外两台中继器,形成单一的闭合通路,所有的工作站都可通过中继器连接到环路上。任何一个工作站发送的信号都可以沿着通信介质进行传播,而且能被所有其他的工作站接收。中继器为环形网提供了 3 种基本功能:数据发送到环中、接收数据和从环中删除数据。它能够接收一个链路上的数据,并以同样的速度串行地把该数据送到另一条链路上,即不在中继器中缓冲。由通信介质及中继器所构成的通信链路是单向的,即能在一个方向上传输数据,而且所有的链路是单向的,即能在一个方向上围绕着环进行循环。

环形拓扑结构的交换方式采用分组交换。由于多个工作站共享同一环,因此需要对此进行控制,以便决定每个工作站在什么时候可以把分组放在环上。一般情况下,环形拓扑结构网络采用令牌环(Token Ring)的介质访问控制。信息发送的过程为:如果某一站点希望将报文发送到另一目的站点,那么它需要将这个报文分成若干个分组。每个分组包括一段数据再加上一些控制信息,其中控制信息包括目的站点的地址。发送信息的站点依次把每个分组放到环上之后,通过其他中继器进行循环;环中的所有中继器都将分组的地址与该中继器连接的节点的地址相比较,当地址符合时,该站点就接收该分组。

环形拓扑结构具有以下优点。

(1) 电缆长度短。环形拓扑结构所需的电缆长度与总线型相当,但比星形要短。

(2) 适用于光纤。光纤传输速率高,环形拓扑网络是单向传输,十分适用于光纤通信介

质。如果在环形拓扑网络中把光纤作为通信介质，将大大提高网络的速度和加强抗干扰的能力。

（3）无差错传输。由于采用点到点通信链路，被传输的信号在每一个节点上再生，因此，传输信息误码率可减到最少。

环形拓扑结构的缺点如下。

（1）可靠性差。在环上传输数据是通过接在环上的每台中继器完成的，所以任何两个节点间的电缆或者中继器故障都会导致全网故障。

（2）故障诊断困难。因为环上的任意一点出现故障都会引起全网的故障，所以难以对故障进行定位。

（3）调整网络比较困难。要调整网络中的配置，例如，扩大或缩小，都是比较困难的。

课后习题

1. 术语解释

计算机网络　网络拓扑结构　局域网　城域网　广域网　通信子网　资源子网

2. 计算机网络的发展可以划分为几个阶段？每个阶段各有什么特点？

3. 以一个你所熟悉的因特网应用为例，说明你对计算机网络定义和功能的理解。

4. 计算机网络如何分类？请分别举出一个局域网、城域网和广域网的实例，并说明它们之间的区别。

5. 计算机网络的二级子网结构如何划分？请说明它们的功能和组成。

6. 常用计算机网络的拓扑结构有哪几种？各自有何特点？试画出它们的拓扑结构图。

7. 计算机网络具有哪些功能？

8. 目前，计算机网络应用在哪些方面？

第2章 网络体系结构与网络协议

学习目的

本章将从层次、服务与协议的基本概念出发，对 OSI 参考模型、TCP/IP 协议与参考模型进行讨论和比较，要求学生掌握网络技术中两个最基本的概念：网络体系结构和网络协议，以便学生对计算机网络的工作原理和实现技术建立一个整体的概念，为以后各个章节的学习打下基础，同时对网络协议标准化和制定国际标准的组织做了系统的介绍。

学习要求

掌握：协议、层次、接口与网络体系结构的基本概念。

掌握：网络体系结构的层次化研究方法。

掌握：OSI 参考模型及各层的基本服务功能。

掌握：TCP/IP 参考模型的层次划分、各层的基本服务功能及主要协议。

理解：OSI 参考模型与 TCP/IP 参考模型比较。

了解：网络协议标准化组织以及 RFC 文档与 Internet 协议标准的制定过程。

2.1 计算机网络体系结构

2.1.1 划分层次结构模型的必要性

为了能够使不同地理分布且功能相对独立的计算机之间组成网络实现资源共享，计算机网络系统需要涉及和解决许多复杂的问题，包括信号传输、差错控制、寻址、数据交换和提供用户接口等一系列问题。计算机网络体系结构是为简化这些问题的研究、设计与实现而抽象出来的一种结构模型。结构模型有多种，如平面结构模型、层次结构模型和网状结构模型等，对于复杂的计算机网络系统，一般采用层次结构模型。在层次结构模型中，往往将系统所要实现的复杂功能分化为若干个相对简单的细小功能，每一项分功能以相对独立的方式去实现。这样，就有助于将复杂的问题简化为若干个相对简单的问题，从而达到分而治之、各个击破的目的。

对网络进行层次划分就是将网络这个庞大的、复杂的问题划分成若干个较小的、简单的问题，即“分而治之”。组成网络部件的组合方式常被描述成它的“体系结构”。而“计算机网络体系结构”采用分层配对结构，定义和描述了一组用于计算机及其通信设施之间互联的标准和规范的集合。

计算机网络层次划分的原则是层内功能内聚，层间耦合松散。也就是说，在网络中，功能相似或紧密相关的模块应放置在同一层；层与层之间应保持松散的耦合，使信息在层与

层之间的流动减小到最小。

计算机网络采用层次化结构的优越性主要体现在以下几个方面。

(1) 各层之间相互独立。高层并不需要知道低层是如何实现的,而仅需要知道该层通过层间的接口所提供的服务。

(2) 灵活性好。当任何一层发生变化时,只要接口保持不变,则在这层以上或以下各层均不受影响。另外,当某层提供的服务不再需要时,甚至可将这层取消。

(3) 各层都可以采用最合适的技术来实现,各层实现技术的改变不影响其他层。

(4) 易于实现和维护。整个系统已被分解为若干个易于处理的部分,这种结构使得一个庞大而又复杂系统的实现和维护变得容易控制。

(5) 有利于网络标准化。因为每一层的功能和所提供的服务都已有了精确的说明,所以标准化变得较为容易。

2.1.2 计算机网络的分层模型

上述分层的思想或方法运用于计算机网络中,就产生了计算机网络的分层模型。在实施网络分层时要依据以下原则。

(1) 根据功能进行抽象分层,每个层次所要实现的功能或服务均有明确的规定。

(2) 每层功能的选择应有利于标准化。

(3) 不同的系统分成相同的层次,对等层次具有相同功能。

(4) 高层使用下层提供的服务时,下层服务的实现是不可见的。

(5) 层的数目要适当,层次太少功能不明确,层次太多体系结构过于庞大。

图 2.1 给出了计算机网络的分层模型示意图,该模型将计算机网络中的每台机器抽象为若干层(Layer),每层实现一种相对独立的功能。分层模型涉及下面一些重要的术语。

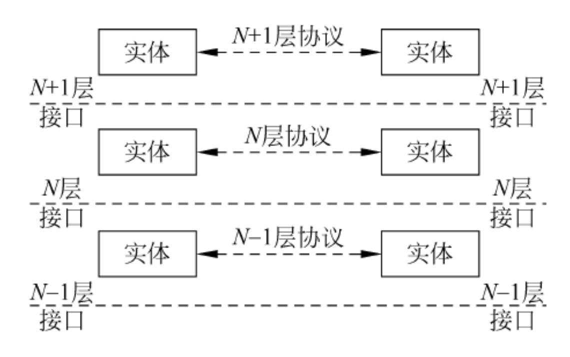


图 2.1 计算机网络的分层模型示意图

1. 实体与对等实体

每一层中,用于实现该层功能的活动元素被称为实体(Entity),包括该层上实际存在的所有硬件与软件,如终端、电子邮件系统、应用程序、进程等。不同机器上位于同一层次、完成相同功能的实体被称为对等(Peer to Peer)实体。

2. 协议

为了使两个对等实体之间能够有效地通信,对等实体需要针对交换什么信息、如何交换信息等问题制定相应的规则或进行某种约定,这种对等实体之间交换数据或通信时所必须

遵守的规则或标准的集合称为协议(Protocol)。

协议由语法、语义和语序三大要素构成。语法包括数据格式、信号电平等；语义是指协议语法成分的含义，包括协调用的控制信息和差错管理；语序包括时序控制和速度匹配关系。

3. 服务与接口

在网络分层结构模型中，每一层为相邻的上一层所提供的功能称为服务。N 层使用 $N-1$ 层所提供的服务，向 $N+1$ 层提供功能更强大的服务。N 层使用 $N-1$ 层所提供的服务时并不需要知道 $N-1$ 层所提供的服务是如何实现的，而只需知道下一层可以为自己提供什么么样的服务，以及通过什么形式提供。N 层向 $N+1$ 层提供的服务通过 N 层和 $N+1$ 层之间的接口来实现。接口定义下一层向其相邻的上一层提供的服务及原语操作，并使下一层服务的实现细节对上一层是透明的。服务是在服务访问点(SAP)提供给上层使用的。N 层 SAP 就是 $N+1$ 层可以访问 N 层服务的地方。每个 SAP 都有一个能够唯一标识它的地址。SDU 是将要跨越网络传递给远方对等实体，然后上交给远方 $N+1$ 层的信息，如图 2.2 所示。

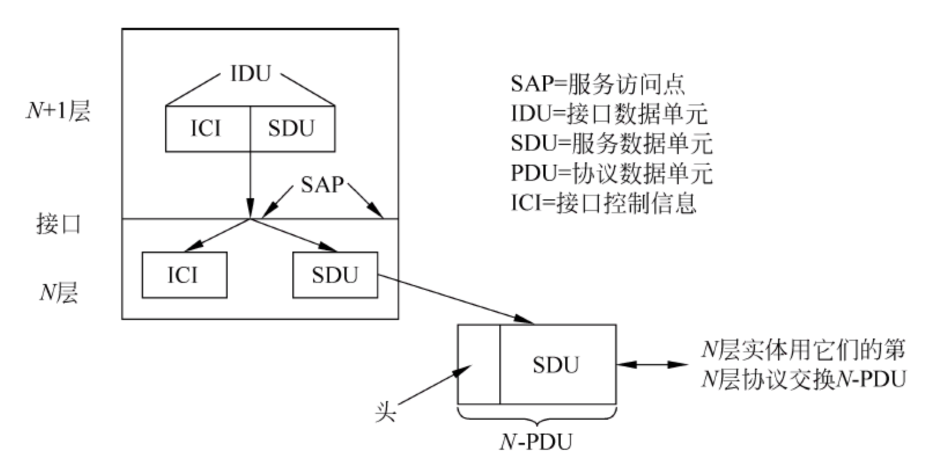


图 2.2 接口层上下层之间的关系

4. 服务类型

在计算机网络协议的层次结构中，层与层之间具有服务与被服务的单向依赖关系，下层向上层提供服务，而上层调用下层的服务。因此可称任意相邻两层的下层为服务提供者，上层为服务调用者。下层为上层提供的服务可分为两类：面向连接服务(Connection Oriented Service)和无连接服务(Connectionless Service)。

(1) 面向连接服务：面向连接服务以电话系统为模式。要和某个人通话，先拿起电话，拨号码，通话，然后挂断。同样在使用面向连接服务时，用户首先要建立连接，使用连接，然后释放连接。连接本质上像个管道：发送者在管道的一端放入物体，接收者在另一端按同样的次序取出物体。其特点是收发的数据不但顺序一致，而且内容也相同。

(2) 无连接服务：无连接服务以邮政系统为模式。每个报文(信件)带有完整的目的地址，并且每一个报文都独立于其他报文，由系统选定的路线传递。在正常情况下，当两个报文发往同一目的地时，先发的先到。但是，也有可能先发的报文在途中延误了，后发的报文反而先到。而这种情况在面向连接服务中是不可能发生的。

5. 服务原语

相邻层之间通过一组服务原语(Service Primitive)建立相互作用,完成服务与被服务的过程,这些原语供用户和其他实体访问该服务。这些原语通知服务提供者采取某些行动或报告某个对等实体的活动。服务原语可被划分为4类,分别是请求(Request)、指示(Indication)、响应(Response)、确认(Confirm)。由不同层发出的每条原语各完成确定的功能,如表2.1所示。

表2.1 4类服务原语

原　语	功　　能(含义)
请　求	服务调用者请求服务提供者提供某种服务
指　示	服务提供者告知服务调用者某事件发生
响　应	服务调用者通知服务提供者响应某事件
确　认	服务提供者告知服务调用者关于它的请求的答复

实体发出连接请求(Connect.request)以后,一个分组就被发送出去。接收方就收到一个连接指示(Connect.indication),被告知一个实体希望和它建立连接。收到连接指示的实体就使用连接响应(Connect.response)原语表示它是否愿意建立连接。但无论是哪一种情况,请求建立连接的一方都可以通过接收连接确认(Connect.confirm)原语获知接收方的态度。

为了更好地理解分层模型及实体、协议、服务、接口等概念,我们以如图2.3所示的邮政系统为例说明这个问题。假设处于A地的用户A要给处于B地的用户B发送信件,为了实现这个信件传递过程,需要涉及用户、邮局和运输部门三个层次。用户A写好信的内容后,将它装在信封里并投入邮筒里交由邮局A寄发;邮局A收到信后,首先进行信件的分拣和整理,然后装入一个统一的邮包交付A地运输部门进行运输,如航空信交民航部门,平信交铁路或公路运输部门等;B地相应的运输部门得到装有该信件的货物箱后,将邮包从其中取出,并交给B地的邮局,B地的邮局将信件从邮包中取出投到用户的信箱中,从而用户B收到了来自用户A的信件。

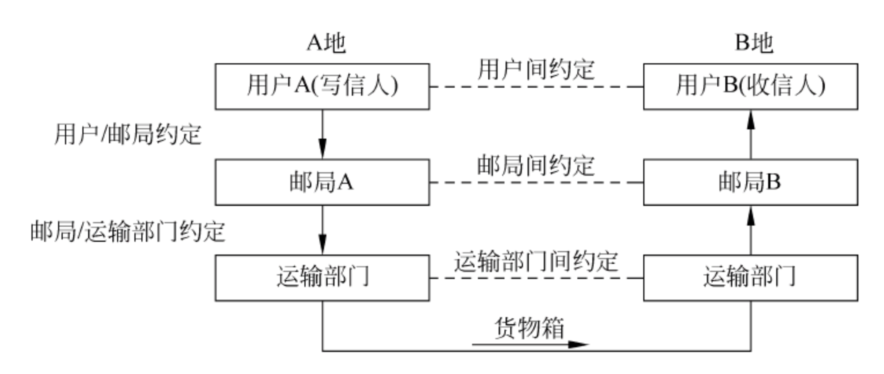


图2.3 网络分层模型的类比——邮政系统模型

在这个过程中,写信人和收信人都是最终用户,处于整个邮政系统的最高层;邮局处于用户的下一层,是为用户服务的。对于用户来说,他只需知道如何按邮局的规定将信件内容装入标准信封并投入邮局设置的邮筒就行了,而无须知道邮局是如何实现寄信过程的,这个过程对用户来说是透明的。处于整个邮政系统最底层的运输部门是为邮局服务的,并且负

责实际的邮件的运送，邮局只需将装有信件的邮包送到运输部门的货物运输接收窗口，而无须操心邮包作为货物是如何到达异地的。在这个例子中，邮筒就相当于邮局为用户提供服务的接口，而运输部门的货物运输接收窗口则是运输部门为邮局提供服务的接口。

另外，在邮政系统的例子中，写信人与收信人、本地邮局和远地邮局、本地运输部门和远地运输部门之间分别还构成了邮政系统分层模型中不同层上的对等实体。为了能将信件准确地由发信人送达收信人，这些对等实体之间必须有一些约定或惯例。例如，写信人写信时必须采用双方都懂的语言文字和文体，开头是对方称谓，最后是落款等。这样，收信人在收到信后才可以读懂信的内容，知道是谁写的，什么时候写的等。同样地，邮局之间要就邮戳的加盖，邮包大小、颜色等制定统一的规则，而运输部门之间也会就货物运输制定有关的航运规定。这些对等实体之间的规则或约定就相当于网络分层模型中的协议。

为了使读者能够更好地理解相关概念，再举一个例子说明。

在日常生活中，两个不同国家的外交官讨论外交事务，可是两人不懂对方的语言，那么他们在讨论外交事务时怎样才能让对方听懂呢？

像这种情况，一般可以请两个翻译：翻译A和翻译B，他们具有以下特点：翻译A会讲英语、法语，翻译B会讲法语、汉语，如图2.4所示，外交官甲将要讨论的外交事务通过英语告诉翻译A，翻译A将英语翻译成法语（法语在这里成了中间语言）再传给翻译B，翻译B将法语翻译成汉语再传给外交官乙，外交官乙就明白了外交官甲的意思。

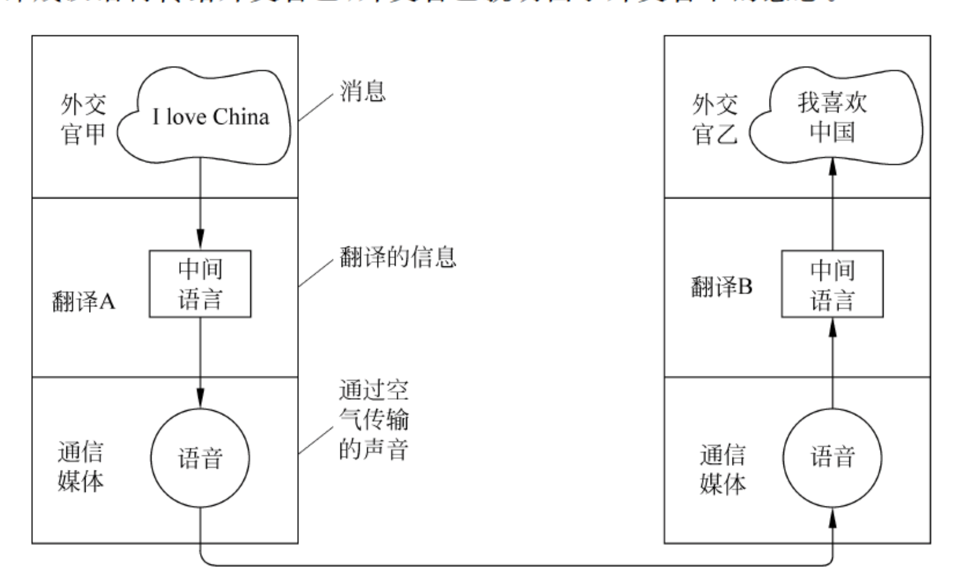


图2.4　外交官问题

通过例子可以知道，翻译们只需翻译语言，并不需要知道外交领域的事务，同样外交官也不需要懂语言翻译领域的事。来自两个不同国家的外交官，他们各自熟悉自己的语言体系，不同的语言有不同的语义和语法规则，他们交流就需要通过掌握多种语言体系的翻译。

计算机网络就是由不同系统的计算机组成的，这些计算机的软硬件环境可能存在较大的差异，就如同不同系统的计算机之间要进行通信也必须通过“翻译”来完成。计算机中的“翻译”就是“网络协议”。

从上面所举可以看出：协议是“水平的”，是控制对等实体间通信的规则；服务是“垂直

的”,是通过层间接口由下层向上层提供的。

从上述关于邮政系统的类比中还可以发现,尽管对收信人来说,信是似乎直接来自写信人,但实际上这封信在 A 地历经了由用户→邮局→运输部门的过程,在 B 地则历经了从运输部门→邮局→用户的过程。

2.1.3 计算机网络体系结构

我们将计算机网络的各层以及其协议的结合,称为网络体系结构。换言之,计算机网络体系结构即指这个计算机网络及其部件所应该完成功能的精确定义。需要强调的是,这些功能究竟由何种硬件或软件完成,则是一个遵循这种体系结构的实现问题。可见体系结构是抽象的,而实现是具体的,是运行在计算机软件和硬件之上的。

网络体系结构是从体系结构的角度来研究和设计计算机网络体系的,其核心是网络系统的逻辑结构和功能分配定义,即描述实现不同计算机系统之间互联和通信的方法与结构,是层和协议的集合。通常采用结构化设计方法,将计算机网络系统划分成若干功能模块,形成层次分明的网络体系结构。

世界上第一个网络体系结构是美国 IBM 公司于 1974 年提出的,它取名为系统网络体系结构(System Network Architecture,SNA)。凡是遵循 SNA 的设备就称为 SNA 设备。这些 SNA 设备可以很方便地进行互联。在此之后,很多公司也纷纷建立自己的网络体系结构,这些体系结构大同小异,都采用了层次技术,但各有其特点以适合本公司生产的计算机组成网络,这些体系结构也有其特殊的名称。如 20 世纪 70 年代末由美国数字网络设备公司 DEC 公司发布的 DNA(Digital Network Architecture,数字网络体系结构)等。但使用不同体系结构的厂家设备是不可以相互连接的。

20 世纪 70 年代末至 80 年代初,一方面是计算机网络规模与数量的急剧增长;另一方面是许多按不同体系结构实现的网络产品之间难以进行相互操作,严重阻碍了计算机网络及各项技术的发展。因此,关于计算机网络体系结构的标准化工作提上了有关国际标准组织的议事日程。

2.2 开放系统互联参考模型

国际标准化组织 ISO 在 1977 年建立了一个分委员会来专门研究体系结构,提出了开放系统互联(Open System Interconnection,OSI)参考模型,这是一个定义连接异种计算机标准的主体结构,并非一个具体的计算机设备或网络,OSI 解决了已有协议在广域网和高通信负载方面存在的问题。

“开放”表示能使任何两个遵守参考模型和有关标准的系统进行连接,这正是“开放”的实际意义。

“互联”是指将不同的系统相互连接起来,以达到相互交换信息、共享资源、分布应用和分布处理的目的。

2.2.1　ISO/OSI 参考模型

ISO/OSI 参考模型的逻辑结构如图 2.5 所示，它将网络结构划分为七层：物理层、数据链路层、网络层、传输层、会话层、表示层和应用层。最低 3 层（1～3）是依赖网络的，涉及将两台通信计算机连接在一起所使用的数据通信网的相关协议，实现通信子网功能。高 3 层（5～7）是面向应用的，涉及允许两个终端用户应用进程交互作用的协议，通常是由本地操作系统提供的一套服务，实现资源子网功能。中间的传输层为面向应用的上 3 层遮蔽了跟网络有关的下 3 层的详细操作，如图 2.6 所示。从实质上讲，传输层建立在由下 3 层提供服务的基础上，为面向应用的高层提供网络无关的信息交换服务。

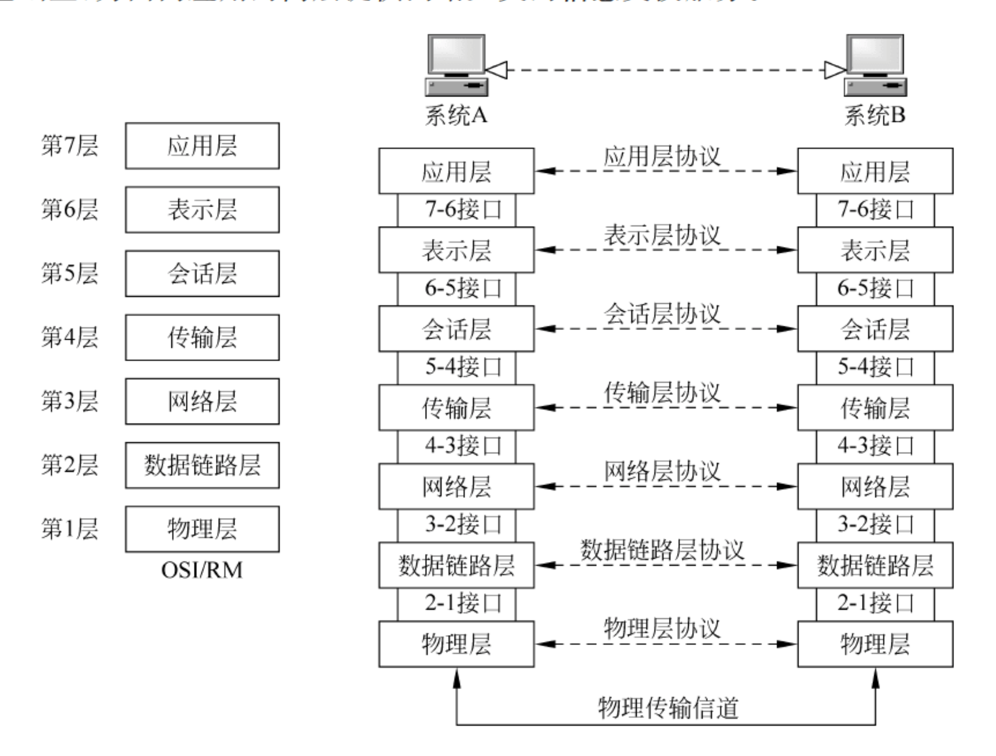


图 2.5　ISO/OSI 参考模型分层结构

每一层均有自己的一套功能集，并与紧邻的上层和下层交互作用。在顶层，应用层与用户使用的软件（如字处理程序或电子表格程序）进行交互。在 OSI 参考模型的底端是携带信号的网络电缆和连接器。总的来说，在顶端与底端之间的每一层均能确保数据以一种可读、无错、排序正确的格式可靠地被发送、接收。

2.2.2　OSI 参考模型各层的功能

OSI 参考模型的每一层都有它自己必须实现的一系列功能，以保证数据报能从源传输到目的地。下面简单介绍 OSI 参考模型各层的功能。

1. 物理层

物理层是 OSI 参考模型的最底层，负责通过物理连接传输比特流，它为数据链路层提供建立、维护和取消物理连接以及在相连的网络系统间传输比特流这两种服务。

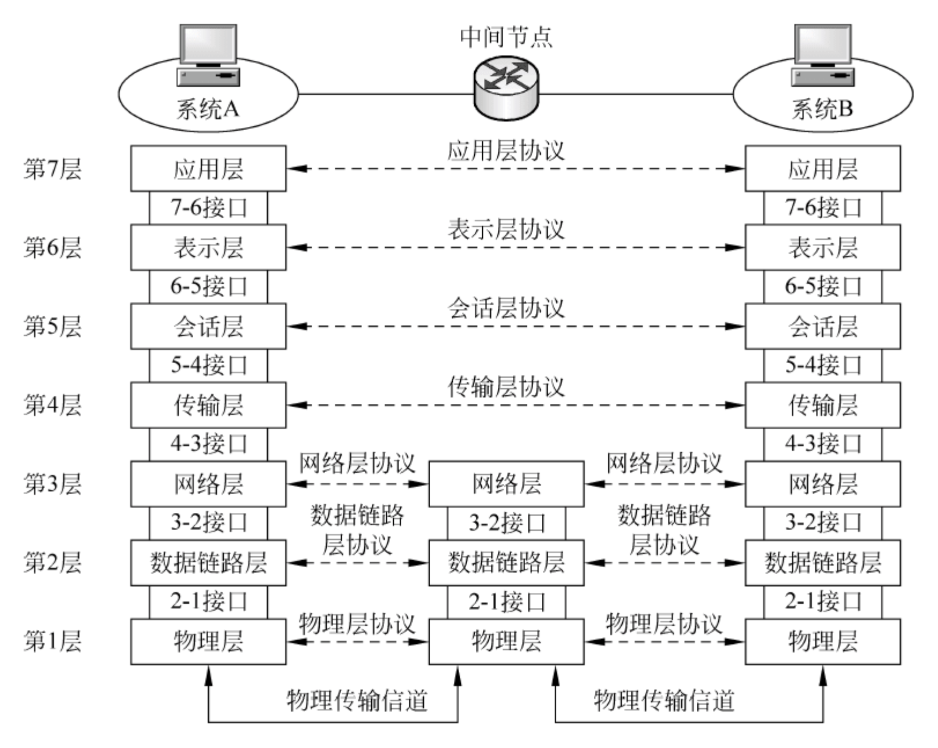


图 2.6 基于 OSI 通信模型结构

物理层定义了网络的物理结构(拓扑)和传输介质的电气、机械规格等有关物理特性。除了不同的传输介质自身的物理特性外,物理层还对通信设备和传输媒体之间使用的接口作了详细的规定。物理层的 4 个特性说明如下。

(1) 机械特性。机械特性规定了物理连接所需接插件的规格尺寸、针脚数量和排列情况等。如 EIA RS-232C 标准规定的 D 型 25 针接口,ITU-T X.21 标准规定的 15 针接口等。

(2) 电气特性。电气特性规定了在物理信道上传输比特流时信号电平的大小、数据的编码方式、阻抗大小、传输速率和距离限制等,比如双绞线不大于 100m,RS-232 接口传输距离不大于 15m,最大速率为 19.2Kbps。

(3) 功能特性。功能特性定义了各个信号线的确切含义,即各个信号线的功能。比如,双绞线每根都有自己的作用。

(4) 规程特性。规程特性定义了利用信号线进行比特流传输的一组操作规程,是指在物理连接的建立、维护和交换信息时数据通信设备之间交换数据顺序。

物理层的协议产生并检测电压以便发送和接收携带数据的信号。在你的桌面 PC 上插入网络接口卡,你就建立了计算机联网的基础。换言之,你提供了一个物理层。尽管物理层不提供纠错服务,但它能够设定数据传输速率并监测数据出错率。网络物理问题,如电线断开,将影响物理层。同样地,如果你没有将网络接口卡在计算机的电路板中插得足够深,计算机也将在物理层出现网络问题。

2. 数据链路层

数据链路层在物理层和网络层之间提供通信,建立相邻节点之间的数据链路,传送按一定格式组织起来的位组合,即数据帧。

数据链路层为网络层提供可靠的信息传送机制。将数据组成适合于正确传输的帧形式。在帧中包含应答、流控制和差错控制等信息,以实现应答、差错控制、数据流控制和发送

顺序控制,确保接收数据的顺序与原发送顺序相同等功能。

3. 网络层

网络层,即 OSI 参考模型的第 3 层,它提供不同网络系统间的连接和路由选择并定义了逻辑地址。该层的数据单元叫作数据包或分组(Packet)。所谓路由就是将数据包从一个网段转发到另一个网段的最佳路径。

为了实现路径选择,网络层必须使用寻址方案来确定存在哪些网络以及设备在这些网络中所处的位置,不同网络层协议所采用的寻址方案是不同的。在确定了目标节点的位置后,网络层还要负责引导数据包正确地通过网络,找到通过网络的最优路径,即路由选择。如果子网中同时出现过多的分组,它们将相互拥塞通路并可能形成网络瓶颈,所以网络层还要提供拥塞控制机制以避免此类现象的出现。另外,网络层还要解决异构网络互联问题。

4. 传输层

数据在传输层进行数据分割和数据重组为数据段(Segment),传输层负责准确可靠地将数据从网络一端传到另一端。

传输层以下的 3 层提供的数据传输有时是不可靠的,传输层加强数据的传输服务,可以将下 3 层的无连接或不受保护的通信升级为面向连接的受保护的通信。

在传输层提供两种服务:面向连接服务和无连接服务。

面向连接服务就像打电话。当与人通电话时,需要拿起听筒并拨号,然后开始交谈,最后挂断电话。与此类似,使用面向连接服务时,首先是建立连接,然后使用连接进行数据传输,最后终止连接。面向连接服务能够保证数据准确可靠地传送到目的地。

无连接服务就像我们寄信一样,我们填写收信人地址和邮政编码并封装好信件后,把它送到邮筒,发信人便完成了通信过程,而信件通过邮局和运输系统最终到达收信人的过程与发信人完全无关。而且,发信人在同时刻发往同一收信人的不同信件,可能会出现晚发的早到情况。所以,在无连接下,当两条消息发送到同一个目的地时,就有可能先发的被延期而后发的先到。但在面向连接服务下,这是不可能发生的。

常见工作在传输层的一种可靠的、面向连接的服务是 TCP/IP 协议套中的 TCP(Transmission Control Protocol,传输控制协议),另一项传输层的服务是 UDP(User Datagram Protocol,用户数据报协议),它是一种不可靠、非面向连接的协议。

同时,传输层还负责在不同物理节点的应用程序间建立连接。因为可能在一个给定的节点上有许多应用程序,它们在同一时间内都在进行通信。比如,在一台计算机上用户有可能在一边收发邮件,一边上网浏览。此时传输层必须使用一种机制来处理节点上的应用程序寻址,使得各个应用程序之间的数据区分开来。传输层使用端口号(Port)明确标识由哪个应用程序处理这些数据。

5. 会话层

会话类似于人们之间的一次谈话。为了使谈话双方能够有序地、完整地进行信息交流,在两个节点间建立、维护和释放面向用户的连接。它在传输层连接的基础上建立会话连接,并进行数据交换管理,允许数据进行单工、半双工和全双工的传送。会话层提供了令牌管理和同步两种服务功能。

6. 表示层

表示层保证一个系统的应用层送出的信息可被另一个系统的应用层所读取，如同应用程序和网络之间的翻译官。如果必要，表示层会利用一种公用的信息表示格式翻译多种信息表示，关心的是所传输数据的语法和语义。

表示层提供的关于数据表示方式的服务有数据表示、数据安全和数据压缩。

7. 应用层

应用层是 OSI 参考模型的第 7 层即最高层，也是最接近使用者的一层。它是计算机网络与最终用户间的接口，它包含了系统管理员管理网络服务所涉及的所有问题和基本功能。它在第六层提供的数据传输和数据表示等各种服务的基础上，为网络用户或应用程序提供完成特定网络服务功能所需的各种应用层协议。简单一点描述应用层应该是：用户通过应用层的协议去完成用户想要完成的任务。

例子：如想上网，那么你会首先打开 IE 浏览器，输入想要冲浪的网址：http://www.cisco.com，如果可以上网会自动出现网页画面，网页本身没有在本地，那怎么可以浏览网页呢，这是因为有了应用层的协议 HTTP(超文本传输协议)来帮助用户与远端的 Web 服务器进行连接且请求传输文件，这样用户就可以通过应用层的协议来完成用户要浏览网页的任务了。

常用的网络服务包括文件服务、电子邮件服务、打印服务、集成通信服务、目录服务、域名解析服务、网络管理、安全和路由互联服务等，如果想要完成类似这样的网络服务都必须通过应用层的协议来完成。

2.2.3 OSI 的层间通信

在同一台计算机的层间交互过程，以及在同一层上不同计算机之间的相互通信过程是相互关联的。

(1) 每一层向其协议规范中的上层提供服务。

(2) 每一层都与其他计算机中相同层的软件和硬件交换一些信息。

1. 同一台计算机之间相邻层的通信

OSI 参考模型描述了在不同计算机上应用程序的信息是如何通过网络介质传输的。对于一个给定的系统的各个层，当要发送的信息逐层向下传输时，信息越往低层就越不同于人类的语言，而是计算机能够理解的 1 和 0。

为了向相邻的高层提供服务，每一层必须知道两层之间定义的标准接口。为了使 N 层获得服务，这些接口定义 $N+1$ 层应该向 N 层提供哪些信息，以及 N 层应向 $N+1$ 层提供何种返回信息。

图 2.7 所示是 OSI 参考模型通信的一个例子。主机 A 发送信息给主机 B。主机 A 的应用程序与主机 B 的应用层通信，主机 A 的应用层再与主机 A 的表示层通信，主机 A 的表示层再与主机 A 的会话层通信，等等，直到到达主机 A 的物理层。物理层把信息放到网络物理介质上并把信息从网络物理介质上送走。信息在网络物理介质上传输并被主机 B 接收后，会以相反的方向向上通过主机 B 的各层(先是物理层，然后是数据链路层，等等)，直到最终到达主机 B 的应用层。

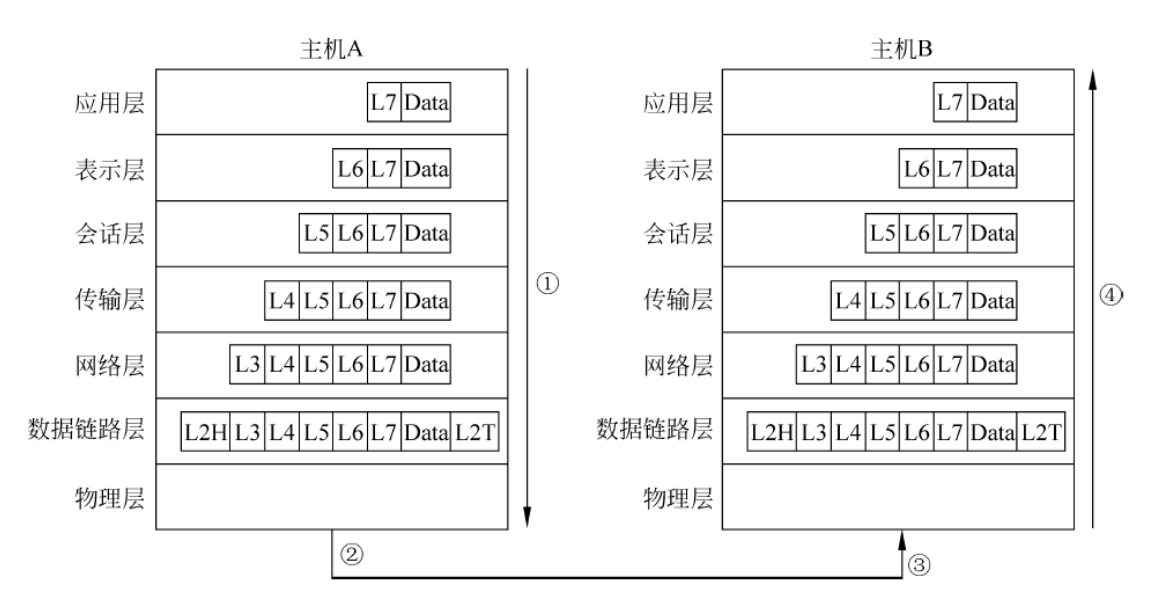


图 2.7 OSI 参考模型相邻层之间通信

注：L＃H—第＃层的头；L＃T—第＃层的尾

数据是由主机 A 中的一些应用程序生成的。例如，用户输入一条 E-mail 消息。每层生成一个头部及所传数据一并传到下一层（图 2.7 中步骤①的箭头，表示数据在不同层的传递过程）。将数据传到下一层意味着下一层需要为上一层提供某种服务。要完成这些服务，下一层需要在包头或包尾中加入一些信息。例如，传输层发送其数据和包头；网络层在其包头中加入正确的网络层目的地址，以使包能被传输到其他计算机上。

从各层的观点来看，在该层包头之后的比特被认为是数据。例如，第 4 层认为第 5 层、第 6 层和第 7 层的包头与原始的用户数据一起是一个大的数据字段。

在应用程序生成数据之后，实现每层功能的软硬件完成各自的工作，加入适当的包头和包尾。为实现物理传输，物理层能够实现媒体发送信号，如图 2.7 中的步骤②。当接收时（步骤③），主机 B 启动其上的相邻层协议进行通信，如图 2.7 中的步骤④，指明了接收数据在协议栈中逐层向上递交处理的过程，具体步骤如下。

（1）物理层（第 1 层）保证比特的同步，并将接收的二进制数据放到缓存中。在将接收到的信号解码成比特流后，通知数据链路层已经收到一个帧。因此，第 1 层在媒体上已经提供了传输的比特流。

（2）数据链路层（第 2 层）检查帧尾的帧校验序列（FCS），判断传输过程中是否有错误发生（差错控制）。如果有错误发生，丢弃此帧。检查数据链路层的地址，使主机 B 决定是否需要进一步处理这些数据。如果这个地址是主机 B 的地址，那么将在第 2 层的包头和包尾之间的数据传递给第 3 层的软件。从而，数据链路层通过该链路实现了数据的传输。

（3）检查网络层（第 3 层）的目的地址。如果该地址是主机 B 的地址（逻辑地址），处理过程将会继续进行，将在第 3 层包头之后的数据传递给传输层（第 4 层）的软件。从而，第 3 层实现了端到端的数据传输服务。

（4）如果传输层（第 4 层）选择了差错恢复，标识这段数据的计数器与确认信息（差错恢复）一起在第 4 层的包头中进行编码。在差错恢复和对输入数据进行重新排序后，将这些数据传递给会话层。

(5) 会话层(第5层)可以用来保证一系列消息的完整性。如果没有完成后续的通信，收到的数据可能没有任何意义。第5层的包头中包含标识字段意味着是一个不连续数据链的中间流而不是结束流。在会话层保证所有的流都完成后，将在第5层包头之后的数据传递给第6层的软件。

(6) 表示层(第6层)定义并维护数据的格式。例如，如果数据是二进制数据而不是字符数据，包头会指明这一点。接收方并不会用主机B中默认的ASCII字符集转换这些数据。通常，此类包头只包括在初始流中，而不包含在每个被传输的流(数据格式)中。在完成了数据格式的转换后，将数据传递给应用层的软件。

(7) 应用层(第7层)处理最后的包头，然后检查真正的终端用户数据。这个包头指明了主机A与主机B已协商好的应用程序所使用的运行参数，该包头用于交换所有参数值。因此，通常只在应用程序初始化时才发送和接收这个包头。例如，在文件传输时，会相互传递所传输文件的长度和文件格式(应用参数)。

2. 不同计算机上同等层之间的通信

第N层必须与另外一台计算机上的第N层通信才能成功地实现该层的功能。例如，传输层(第4层)能够发送数据，但如果另外一台计算机不对那些已接收的数据进行确认，那么发送方就不知应在何时进行差错恢复。同样，发送方计算机将网络层目的地址(第3层)放到包头中。如果中继路由器拒绝合作，不执行网络层功能，那么包就不会被传输到真正的目的地。

为了与其他计算机上的同等层进行通信，每一层都定义了一个包头，而且有时还定义了包尾。包头和包尾是附加的数据位，由发送方计算机的软件或硬件生成，放在由第$N+1$层传给第N层的数据的前面或后面。这一层与其他计算机上同等层进行通信所需要的信息就在这些包头或包尾中被编码。接收方计算机的第N层软件或硬件解释由发送方计算机第N层所生成的包头或包尾编码，从而得知此时第N层的过程应如何处理。

每一层使用自己层的协议与其他系统的对等层相互通信。每一层的协议与对等层之间交换的信息称为协议数据单元(PDU)。

如图2.8所示提供了同等层之间通信的概念模型。主机A的应用层与主机B的应用层通信。同样，主机A的传输层、会话层和表示层也与主机B的对等层进行通信。OSI参考模型的下3层必须处理数据的传输，路由器C参与此过程。主机A的网络层、数据链路层和物理层与路由器C进行通信。同样，路由器C与主机B的物理层、数据链路层和网络层进行通信。

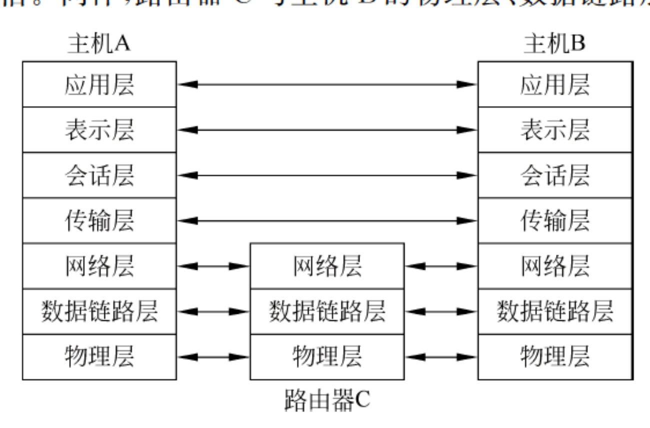


图2.8 OSI参考模型对等层通信

OSI 参考模型的分层禁止了不同主机的对等层之间的直接通信。因此，主机 A 的每一层必须依靠主机 A 相邻层提供的服务来与主机 B 的对等层通信。假定主机 A 的第 4 层必须与主机 B 的第 4 层通信，那么，主机 A 的第 4 层就必须使用主机 A 的第 3 层提供的服务。第 4 层叫作服务用户，第 3 层叫作服务提供者。第 3 层通过一个服务访问点(SAP)给第 4 层提供服务。这些服务访问点使得第 4 层能要求第 3 层提供服务。

3. 封装

通常将数据放置在每一层的包头后面(及包尾之前)的概念称为封装。如图 2.9 所示，当每一层生成了包头时，将由相邻上一层传递来的数据放到该包头的后面，这样就封装了高一层的数据。对数据链路层(第 2 层)协议而言，第 3 层的包头和数据将放到第 2 层的包头和包尾之间。物理层并不使用封装，因为它不使用包头和包尾。参考图 2.9，从用户数据的生成到编码物理信号的整个封装过程如下所述。

(1) 应用程序已经生成了数据。应用层生成该层的包头并将数据放在其后，并将这个数据传递到表示层。

(2) 表示层生成该层的包头并将数据放在其后，这个数据结构被传递到会话层。

(3) 会话层生成该层的包头并将数据放在其后，这个数据结构被传递到传输层。

(4) 传输层生成该层的包头并将数据放在其后，这个数据结构被传递到网络层。

(5) 网络层生成该层的包头并将数据放在其后，这个数据结构被传递到数据链路层。

(6) 数据链路层生成该层的包头并将数据放在其后，数据链路层把包尾放到此结构的后面，这个数据结构被传递到物理层。

(7) 物理层在媒体上对信号进行编码，传输该数据位。

上面的 7 步过程对于 OSI 参考模型是准确、有意义的。然而，对于应用的每次数据传输来说，(通常)并不是每一层都进行封装。一般来说，第 5 层到第 7 层在初始化期间使用包头。但在大多数的流中，没有第 5 层、第 6 层或第 7 层的包头，如图 2.9 所示，这是因为不是每个流的数据都有新的信息需要交换。

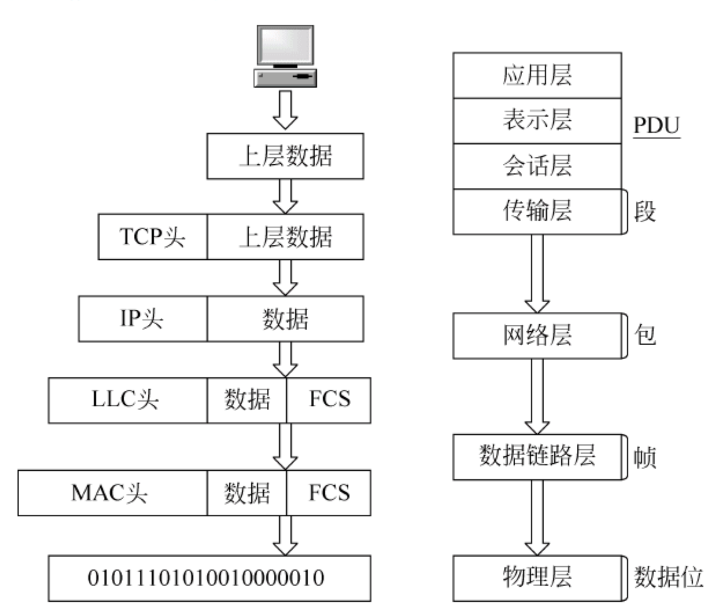


图 2.9　发送方的数据的封装过程

4. 解封装

当远程设备顺序接收到一串比特时，它会把它们传输给数据链路层以组装为帧。当数据链路层接收到该帧时，它会执行以下工作。

(1) 读取物理地址和由直接相连的对等数据链路层所提供的控制信息。

(2) 从该帧剥离该控制信息并由此创建一个数据报。

(3) 遵照在帧的控制部分中出现的内容而把数据报向上传输到相邻层。

这个过程被称为解封装(De-Encapsulation)，每个后续层都会经历一个类似过程，如图 2.10 所示。

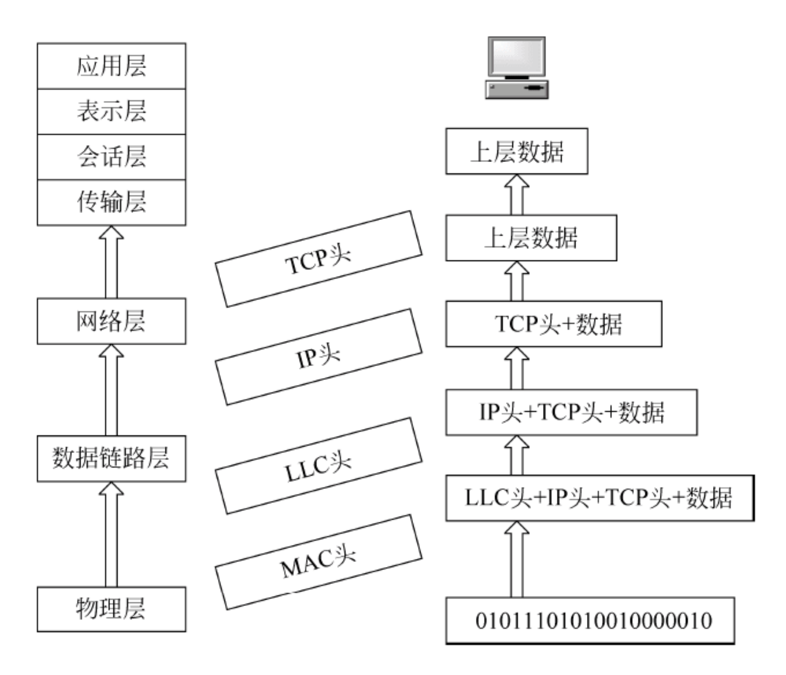


图 2.10 接收方的数据的解封装过程

5. OSI 参考模型每一层数据的名称

为了使数据分组从源主机传输到目的主机，如图 2.11 所示，源主机 OSI 参考模型的每一层要与目的主机的每一层进行通信。用 Peer-to-Peer Communications(对等实体间通信)表示源主机与目的主机对等层间的通信。在这一过程中，每一层的协议交换的信息称为协议数据单元(PDU)。位于源计算机上的每个通信层，使用针对该层的协议数据单元与目标计算机上的对等层进行通信。

数据分组起源于一台源主机，并且被传输到一台目的主机上。每一层需要依赖于其低层提供的服务。为了提供这种服务，低层将来自上层的协议数据单元封装到它的数据字段中，然后增加实现本层功能所需的包头和包尾。之后，当数据通过 OSI 参考模型的每一层时，都会增加相应的包头和包尾。在第 7 层、第 6 层和第 5 层增加了它们的信息后，第 4 层需要增加更多的信息。第 4 层的协议数据单元被称为段(Segment)。

比如，网络层为传输层提供服务，传输层将数据提交给网络层。网络层的任务是将数据通过网络进行传输，它通过对该数据进行封装和增加一个包头，创建一个分组或数据报(第 3 层的 PDU)来实现这一任务。数据报头里包含为完成传输所需的信息，如源和目的逻辑地址。

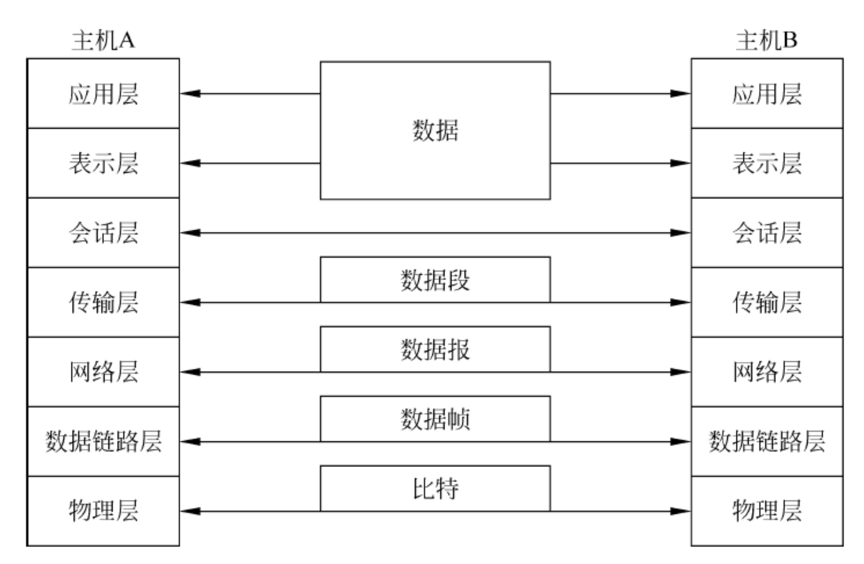


图 2.11　PDU 名称

数据链路层为网络层提供服务，它将网络层信息封装到帧（第 2 层 PDU）中，帧包头包含完成链路操作的信息（如物理地址）。数据链路层通过将网络层数据封装到帧中的方式为网络层提供服务。此外，数据链路层包头信息是本地的，而且只对直接相连的设备有意义。

物理层也为数据链路层提供服务，物理层将数据链路层帧编码成 1 和 0 的模式以便在介质上传输（通常为线缆），然后这些比特被传输到位于端到端路径中的下一台直接相连的设备上。

2.3　TCP/IP 参考模型

尽管 OSI 参考模型得到了全世界的认同，但是互联网历史上和技术上的开发标准都是 TCP/IP（传输控制协议/网际协议）参考模型。

2.3.1　TCP/IP 参考模型概述

TCP/IP 参考模型是由美国国防部创建的，所以有时又称 DoD（Department of Defense）模型，是迄今为止发展最成功的通信协议，它被用于构筑目前最大的、开放的互联网系统 Internet。TCP/IP 是一组通信协议的代名词，这组协议使任何具有网络设备的用户能访问和共享 Internet 上的信息，其中最重要的协议族是传输控制协议（TCP）和网际协议（IP）。

TCP 和 IP 是两个独立且紧密结合的协议，负责管理和引导数据报文在 Internet 上的传输。二者使用专门的报文头定义每个报文的内容。TCP 负责和远程主机的连接，IP 负责寻址，使报文被送到其该去的地方。

TCP/IP 也分为不同的层次开发，每一层负责不同的通信功能。但 TCP/IP 协议简化了层次设备（只有 4 层），由下而上分别为网络接口层、网络层、传输层、应用层，如图 2.12 所示。

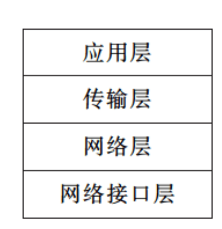


图 2.12　TCP/IP 协议模型

(1) 网络接口层：有时也称为数据链路层，通常包括操作系统中的设备驱动程序和计算机中对应的网络接口卡。它们一起处理与电缆（或其他任何传输媒介）的物理接口细节。

(2) 网络层：有时也称为互联网层，处理分组在网络中的活动，如分组的选路。在TCP/IP 协议族中，网络层协议包括 IP（网际协议）、ICMP（Internet 控制消息协议），以及 IGMP（Internet 组管理协议）。

(3) 传输层：主要为两台主机上的应用程序提供端到端的通信。在 TCP/IP 协议族中，有两个互不相同的传输协议：TCP（传输控制协议）和 UDP（用户数据报协议）。

TCP 为两台主机提供高可靠性的数据通信。它所做的工作包括把应用程序交给它的数据分成合适的小块交给下面的网络层，确认接收到的分组，设置发送最后确认分组的超时时钟等。由于传输层提供了高可靠性的端到端的通信，因此应用层可以忽略所有这些细节。而另一方面，UDP 则为应用层提供一种非常简单的服务。它只是把称为数据报的分组从一台主机发送到另一台主机，但并不保证该数据报能到达另一端。任何必需的可靠性必须由应用层来提供。这两种传输层协议分别在不同的应用程序中有不同的用途，这一点将在后面看到。

(4) 应用层：负责处理特定的应用程序细节。几乎各种不同的 TCP/IP 实现都会提供下面这些通用的应用程序。

① Telnet 远程登录；

② FTP 文件传输协议；

③ SMTP 简单邮件传输协议；

④ SNMP 简单网络管理协议。

2.3.2 各层主要协议

TCP/IP 事实上是一个协议系列或协议族，目前包含了 100 多个协议，用来将各种计算机和数据通信设备组成实际的 TCP/IP 计算机网络。TCP/IP 参考模型各层的一些重要协议如图 2.13 所示。

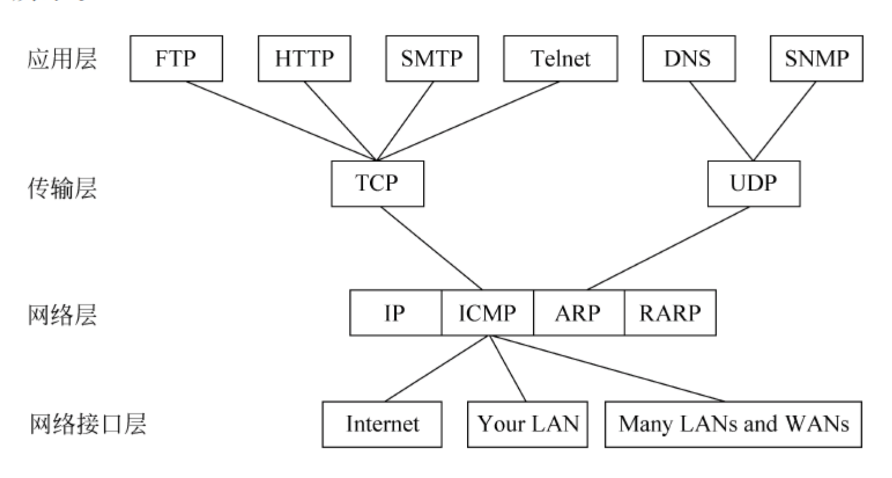


图 2.13 TCP/IP 协议图

1. 网络接口层协议

TCP/IP 的网络接口层中包括各种物理网协议，例如，Ethernet、令牌环、帧中继、ISDN 和分组交换网 X.25 等。当各种物理网被用作传输 IP 数据报的通道时，就可以认为是属于这一层的内容。

2. 网络层协议

网络层包括多个重要协议，主要协议有 4 个，即 IP、ICMP、ARP 和 RARP。

(1) 网际协议(Internet Protocol，IP)是其中的核心协议，IP 协议规定网络层数据分组的格式。

(2) Internet 控制消息协议(Internet Control Message Protocol，ICMP)：提供网络控制和消息传递功能。

(3) 地址解释协议(Address Resolution Protocol，ARP)：用来将逻辑地址解析成物理地址。

(4) 反向地址解释协议(Reverse Address Resolution Protocol，RARP)：通过 RARP 广播，将物理地址解析成逻辑地址。

3. 传输层协议

传输层的主要协议有 TCP 和 UDP。

(1) 传输控制协议(Transport Control Protocol，TCP)：是面向连接的协议，用三次握手和滑动窗口机制来保证传输的可靠性和进行流量控制。

(2) 用户数据报协议(User Datagram Protocol，UDP)：是面向无连接的不可靠传输层协议。

4. 应用层协议

应用层包括了众多的应用与应用支撑协议。常见的应用层协议有文件传输协议(FTP)、超文本传输协议(HTTP)、简单邮件传输协议(SMTP)、远程登录(Telnet)；常见的应用层支撑协议包括域名服务(DNS)和简单网络管理协议(SNMP)等。

2.4 OSI 参考模型和 TCP/IP 参考模型的区别

1. 相似点

ISO/OSI 参考模型和 TCP/IP 参考模型有许多相似之处，具体表现在：两者均采用了层次结构并存在可比的传输层和网络层；两者都有应用层，虽然所提供的服务有所不同；均是一种基于协议数据单元的包交换网络，而且分别作为概念上的模型和事实上的标准，具有同等的重要性。

2. 不同点

ISO/OSI 参考模型和 TCP/IP 参考模型还有如下许多不同之处。

(1) OSI 参考模型包括了 7 层，而 TCP/IP 参考模型只有 4 层。虽然它们具有功能相当的网络层、传输层和应用层，但其他层并不相同。

(2) TCP/IP 参考模型中没有专门的表示层和会话层，它将与这两层相关的表达、编码和会话控制等功能包含到了应用层中去完成。另外，TCP/IP 参考模型还将 OSI 参考模型

的数据链路层和物理层包括到了一个网络接口层中。

(3) OSI参考模型在网络层支持无连接和面向连接两种服务,而在传输层仅支持面向连接服务。TCP/IP参考模型在网络层则只支持无连接一种服务,但在传输层支持面向连接和无连接两种服务。

(4) TCP/IP由于有较少的层次,因而显得更简单,TCP/IP一开始就考虑到多种异构网的互联问题,并将网际协议(IP)作为TCP/IP的重要组成部分,并且作为从Internet上发展起来的协议,已经成了网络互联的事实标准。但是,目前还没有实际网络是建立在OSI参考模型基础上的,OSI仅仅作为理论的参考模型被广泛使用。

课后习题

1. 术语解释

网络体系结构　服务　接口　协议　实体　协议数据单元　数据封装　数据解封装

2. 在OSI参考模型中,保证端—端的可靠性是在(　　)上完成的。

A. 数据链路层　B. 网络层　C. 传输层　D. 会话层

3. 数据的加密和解密属于OSI参考模型(　　)的功能。

A. 网络层　B. 表示层　C. 物理层　D. 数据链路层

4. OSI参考模型包括哪7层?试解释每层的主要功能。

5. 同一台计算机之间相邻层如何通信?

6. 不同计算机上同等层之间如何通信?

7. 简述OSI参考模型各层的功能。

8. 简述数据发送方封装的过程。

9. OSI参考模型中每一层数据单元分别是什么?

10. 在TCP/IP协议中各层有哪些主要协议?

11. 试说明层次、协议、服务和接口的关系。

12. 计算机网络为什么采用层次化的体系结构?试举出一些与分层体系结构的思想相似的日常生活的例子。

13. 试比较TCP/IP参考模型和OSI参考模型的异同点。

14. 长度为100B(字节)的应用层数据交给运输层传送,需加上20B的TCP首部。再交给网络层传送,需加上20B的IP首部。最后交给数据链路层的以太网传送,加上首部和尾部共18B。求数据的传输效率。数据的传输效率是指发送的应用层数据除以所发送的总数据(即应用数据加上各种首部和尾部的额外开销)。若应用层数据长度为1000B,数据的传输效率是多少?

第3章 数据通信基础与物理层

学习目的

数据通信技术是网络技术发展的基础。本章将对数据通信的基本概念、主要的传输介质、数据编码技术、数据传输技术、多路复用技术与差错控制技术进行系统的分析。学习本章内容将对网络中最基本的数据通信技术、广域网中数据传输原理与实现方法的理解有很大的帮助，为以后的学习打下坚实的基础。

学习要求

理解：数据通信的基本概念。

掌握：传输介质类型及主要特性。

掌握：数据编码的类型和基本方法。

掌握：基带传输与频带传输的基本概念。

掌握：数据传输方式的基本概念。

掌握：多路复用的分类与特点。

掌握：宽带接入技术。

3.1 物理层

物理层处于网络参考模型的最底层，它的上一层是数据链路层，它向下直接和传输介质相连。物理层不仅仅指与计算机相连接的物理设备或者具体的传输介质，主要考虑的是如何在连接开放系统的不同传输介质上传输各种数据的比特流。

我们知道，现有计算机网络中的硬件设备和传输媒体的种类非常繁多，通信手段也有许多不同方式，存在着很大的差异；与此同时，各种新的通信技术又在快速发展，因此网络设计中，试图通过设计物理层来尽可能屏蔽这些差异，使数据链路层只需考虑本层的服务和设计，而不需要考虑物理层具体使用了哪些传输介质和物理传输设备。用于物理层的协议也称为物理层规程(Procedure)。只是在“协议”这个名词出现之前人们就先使用了“规程”这一名词。

物理层的主要任务描述为确定与传输媒体接口有关的一些特性，具体如下。

(1) 机械特性：指明接口所用接线器的形状和尺寸、引线数目和排列、固定和锁定装置等。这很像平时常见的各种规格的电源插头的尺寸都有严格的规定。

(2) 电气特性：指明在接口电缆的各条线上出现的电压的范围。

(3) 功能特性：指明某条线上出现的某一电平的电压表示何种意义。

(4) 规程特性：指明对于不同功能的各种可能事件的出现顺序。

数据在计算机中多采用并行传输方式，但数据在通信线路上的传输方式一般都是串行传输（这是出于经济上的考虑），即逐个比特按照时间顺序传输，因此物理层还要完成传输方式的转换。

具体的物理层协议种类繁多，这是因为物理连接的方式很多（如可以使用点对点的，也可以采用多点连接或广播连接），而传输媒体的种类也非常多（双绞线、同轴电缆、光缆以及各种波段的无线信道等）。因此在学习物理层时，应将重点放在掌握基本概念上。

考虑到使用本书的一部分读者可能没有学过有关数据通信方面的课程，因此我们利用下面的3.2节简单地介绍一下有关现代通信的一些最基本的知识和最重要的结论（不给出证明过程）。对于已具有这方面知识的读者可略过这部分内容。

3.2 数据通信

从某种意义上讲，计算机网络是建立在数据通信系统之上的资源共享系统。计算机网络的主要功能是为了实现信息资源的共享与交换，而信息是以数据形式来表达的，所以计算机网络必须解决数据通信的问题。

3.2.1 数据通信的基本概念

1. 信息、数据和信号

信息是指有用的知识或消息，计算机网络通信的目的就是为了交换信息。而数据则是运送信息的实体，是信息的表达方式，可以是数字、文字、声音、图形和图像多种不同形式。在计算机系统中，统一以二进制代码表示数据的不同形式。而当这些二进制代码表示的数据要通过物理介质和器件进行传输时，还需要将其转变成物理信号，信号（Signal）是数据在传输过程中的电磁波表达形式。

2. 模拟信号与数字信号

作为数据的电磁波表达形式，信号一般以时间为自变量，以表示数据的某个参量如振幅、频率或相位为因变量，并且按其因变量对时间的取值是否连续被分为模拟信号和数字信号。模拟信号是指信号的因变量随时间连续变化的信号，如图3.1所示。电视图像信号、语音信号、温度压力传感器的输出信号以及许多遥感遥测信号都是模拟信号。数字信号是指信号的因变量不随时间连续变化的信号，通常表现为离散的脉冲形式，可表示为 $x(nt)$，如图3.2所示。显然，在数字信号中，因变量取值状态是有限的。计算机数据、数字电话和数字电视等都可看成是数字信号。

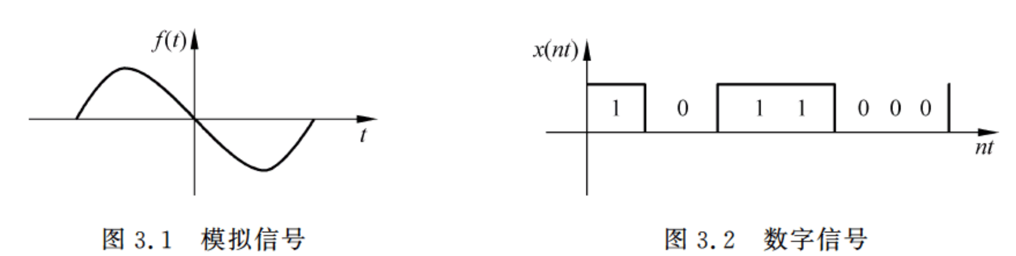


图3.1 模拟信号　　图3.2 数字信号

虽然模拟信号与数字信号有着明显的差别，但二者之间在一定条件下是可以相互转换的，转换可以通过调制解调器来完成。模拟信号可以通过采样、编码等步骤转换成数字信号，而数字信号也可以通过解码、平滑等步骤转换成模拟信号。

3. 数据通信

发送方将要发送的数据转换成信号通过物理信道传输到数据接收方的过程称为数据通信。由于信号可以是离散变化的数字信号，也可以是连续变化的模拟信号，所以与之相对应，数据通信被分为模拟数据通信和数字数据通信。所谓模拟数据通信，是指在模拟信道上以模拟信号形式来传输数据；而数字数据通信则是指利用数字信道以数字信号形式来传输数据。

4. 源点、终点和信道

在数据通信中，通常将数据的发送方称为源点，而将数据的接收方称为终点。源点和终点一般是计算机或其他一些数据终端设备。

为了在源点和终点之间实现有效的数据传输，必须在源点和终点之间建立一条传输信号的物理通道，这条通道被称为物理信道，简称信道。信道建立在传输介质之上，但包括了传输介质和附属的通信设备。通常，同一传输介质上可提供多条信道，一条信道允许一路信号通过。按传输介质的类型来划分，信道被分为有线信道和无线信道；按信道中所传输的信号类型来划分，信道被分为模拟信道和数字信道。

3.2.2　数据通信系统的模型

数据通信系统是指通过通信线路和通信控制处理设备将分布在各处的数据终端设备连接起来，执行数据传输功能的系统。图 3.3 给出了数据通信系统的模型。

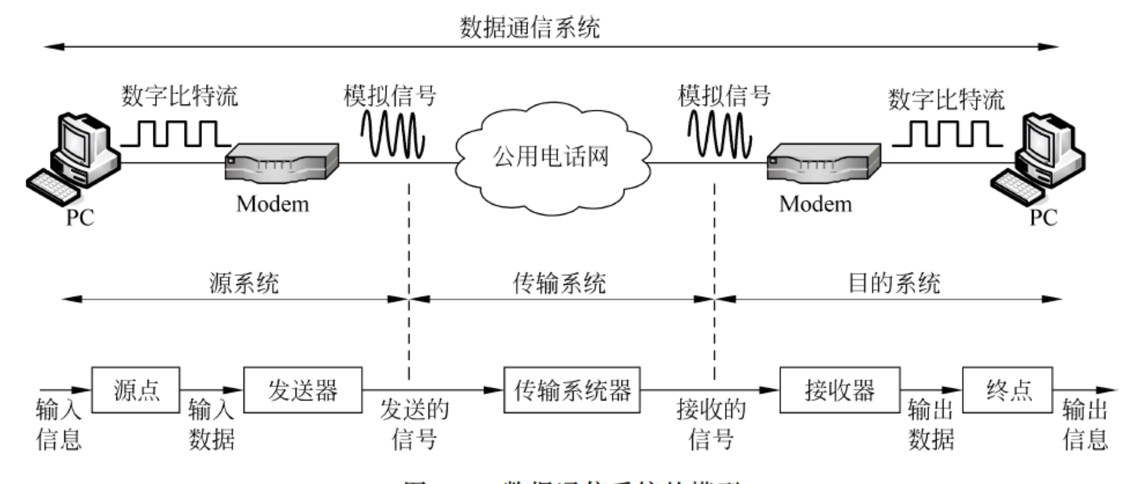


图 3.3　数据通信系统的模型

一个数据通信系统由源系统（或发送端）、传输系统（或传输网络）和目的系统（或接收端）3 个部分组成。

源系统一般包括以下两部分。

（1）源点：源点设备发送要传输的数据，又称源站，发送输出的数字比特流。

（2）发送器：通常源点发送的数据要通过发送器编码后才能够在传输系统中进行传

输。发送器把源点所要发送的数据转换成适合于在信道上传输的信号。

目的系统一般包括以下两部分。

(1) 接收器：接收传输系统传输过来的信号，并将其转换为能够被目的设备处理的信息。把从信道上接收的信号转换成终点所能识别的数据。

(2) 终点：终点设备从接收器获取传输来的数据，又称目的站。

源点和终点分别是数据的出发点与目的地，又称为数据终端设备(Data Terminal Equipment，DTE)。DTE 通常属于资源子网的设备，如资源子网中的计算机、数据输入/输出设备和通信处理机等。

发送器和接收器又称为数据线路端接设备(Data Circuit-terminating Equipment，DCE)。DCE 为 DTE 提供了入网的连接点，通常被认为是通信子网中的设备。

3.2.3 数据通信系统的主要质量指标

衡量和评价一个系统的好坏，必须涉及系统的主要性能指标问题。数据通信的主要质量指标是衡量数据传输的有效性和可靠性的参数。有效性是指传输一定的信息量所消耗的信息资源(带宽或时间)，主要由数据传输的数据速率、调制速度、传输延迟、信道带宽和信道容量等指标来衡量；而可靠性是指接收信息的准确程度，一般由数据传输的误码率指标来衡量。有效性和可靠性这两个要求通常是矛盾的，因此只能根据需要及技术发展水平尽可能取得适当的统一。例如，在一定的可靠性指标下，尽可能提高信息的传输速率；或者在一定有效性条件下，使消息的传输质量尽可能提高。模拟通信和数字通信对这两个指标要求的具体内容有较大差异。

1. 模拟通信系统的质量指标

(1) 有效性。模拟通信系统的有效性是用有效传输带宽来度量，同样的信息采用不同的调制方式，则需要不同的频带宽度。频带宽度越窄，有效性越好。

(2) 可靠性。模拟通信系统的可靠性是用接收端最终的输出信噪比来度量，信噪比越大，通信质量越高。如普通电话要求信噪比在 20dB 以上，电视图像则要求信噪比在 40dB 以上。

2. 数字通信系统的质量指标

数字通信系统中，有效性用传输速率来表示，可靠性用差错率(误码率)来衡量。

信道是通信双方以传输介质为基础的传输信息的通道，它是建立在通信线路及其附属设备(如收发设备)上的。表面上看，信道与传输介质好像差不多，但信道又不能等同于传输介质，同一条传输介质可以同时存在多条信号通道，即一条通信介质构成的线路上往往包含了多条信道。与信号的分类相似，信道也可分为传送模拟信号的模拟信道和传送数字信号的数字信道两大类。但数字信号经过数/模变换后就可以在模拟信道上进行传送，而模拟信号经过模/数变换后也可以在数字信道上进行传送。由有线传输介质(如双绞线、同轴电缆、光缆等)构成的通信信道叫作有线信道；由无线传输介质(如微波、卫星)构成的通信信道叫作无线信道。信道带宽或信道容量是描述信道的主要指标之一，由信道的物理特性所决定。

信道带宽是指信道中能够传送的信号的频率范围。当信号的带宽超过信道带宽时，信号就不能在该信道上传送，或者传送的信号将会失真。为计算带宽，需要在频率范围内用最

高频率减去最低频率。例如，最高频率为 5000Hz，最低频率为 1000Hz，则带宽即为 4000Hz。

信道容量是指单位时间内信道所能传输的最大信息量，即一个信道能够达到的最大的传输速率，它表示信道的传输能力。在通信领域中，信道容量常指信道在单位时间内可传输的最大码元数(码元是承载信息的基本信号单位，一个表示数据有效值状态的脉冲信号就是一个码元，其单位为波特)，信道容量以码元速率(或波特)来表示。由于数据通信主要是计算机与计算机之间的数据传输，而这些数据最终又以二进制位的形式表示，因此，信道容量有时也表示为单位时间最多可传输的二进制的位数(也叫信道的数据传输速率)，以位/秒(b/s)形式表示，简称为 bps。

按信道频率范围的不同，通常可将信道分为 3 类：窄带信道(带宽为 0～300Hz)、音频信道(带宽为 300～3400Hz)和宽带信道(带宽为 3400Hz 以上)。

(1) 数据传输速率(Rate)。数据传输速率是指通信系统单位时间内传送的二进制代码的位数(比特数)，因此又称比特率，单位用比特/秒表示，记为 b/s 或 bps。

数据传输速率的高低，由每位数据所占的时间来决定，一位数据所占的时间宽度越小，其数据传输速率就越高。设 T 为传输的电脉冲的宽度或周期，N 为一个脉冲信号所有可能的状态数，则数据传输速率为

$$R_S = \frac{1}{T}\log_2 N \quad \text{(bps)}$$

式中，$\log_2 N$ 是每个电脉冲信号所表示的二进制数据的位数(比特数)。如电信号的状态数 $N=2$，即只有 0 和 1 两个状态，则每个电信号只传送 1 位二进制数据，此时，$R_S=1/T$。

(2) 调制速率。调制速率又称波特速率或码元速率，它是数字信号经过调制后的传输速率，表示每秒传输的电信号单元(码元)数，即调制后模拟电信号每秒钟的变化次数，它等于调制周期(即时间间隔)的倒数，单位为波特(Baud)。若用 T(秒)表示调制周期，则调制速率为 $R_b=1/T$(Baud)，即 1 波特表示每秒钟传送一个码元。

显然，上述两个指标有如下的数量关系：$R_S=R_b\log_2 N$(bps)，即在数值上“波特”单位等于“比特”的 $\log_2 N$ 倍，只有当 $N=2$(即双值调制)时，两个指标才在数值上相等。但是，在概念上两者并不相同，Baud 是码元的传输速率单位，表示单位时间传送的信号值(码元)个数，波特速率是调制速度，而 bps 是单位时间内传输信息量的单位，表示单位时间传送的二进制数的个数。

误码率是衡量通信系统在正常情况下传输可靠性的指标。误码率是指二进制码元在传输过程中被传错的概率。显然，它就是错误接收的码元数在所传输的总码元数中所占的比例。误码率的计算公式为

$$P_e = N_e/N$$

式中，P_e表示误码率；N_e表示被传错的码元数；N 表示传输的二进制码元总数。上式在 N 取值很大时才有效。在计算机网络通信系统中，要求误码率低于 10^{-6}。如果实际传输的不是二进制码元，需要折合成二进制码元来计算。在通信系统中，系统对误码率的要求应权衡通信的可靠性和有效性。

信道的带宽是由硬件设备改变电信号的跳变响应时间决定的。尽管信号的传输速率为 300 000km/s，但由于发送和接收设备存在响应时间，特别是计算机网络系统中的通信子网

还存在中间转发等待时间，以及计算机系统的发送和接收处理时间，所以，在系统的信息传输过程中存在着延迟(传输延迟)。

在计算机网络中由于不同的通信子网和不同的网络体系结构采用不同的中转控制方式，因此，在通信子网中存在的中转延迟只能依据网络状态而定。由电信号响应带来的延迟时间则是固定的。显然，响应时间越小，延迟就越小。也就是说，信道的带宽越大，延迟就越小。

传输延迟是指由于各种原因的影响，使得系统信息在传输过程中存在着不同程度的延迟或滞后的现象。信息的传输延迟时间包括发送和接收处理时间、电信号响应时间、中间转发时间和信道传输时间等。传输延迟通常又分为传输时延和传播时延。

传输时延：是指发送一组信息所用的时间，该时间与信息传输速率和信息格式有关。

传播时延：是指信号在物理媒体中传输一定距离所用的时间，它与信号传输速率和距离有关。人们都知道，在理想的情况下，电磁波的传输速率为 300 000km/s(即光速)。通常认为电磁波在光纤、卫星信道中的传输速率可达到光速，而在一般电缆中的传输速率约为光速的 2/3。用下面的例题来更好地理解传输时延和传播时延。

例：在相隔 1000km 的两地传输 3Kb 的数据，可以通过电缆以 20Kbps 的速率传输或通过卫星信道以 60Kbps 的速率传输，问从发送方开始到接收方接收到全部数据用哪种方式时间较短？(假定信息在电缆中传输速率为 200 000km/s，而在卫星信道中的传输速率是 300 000km/s，卫星距离地面 36 000km)。

数据在电缆中的传输时延为 3Kb/20Kbps＝150ms，而其传播时延为 1000km/(2×10^5 km/s)＝5ms，因此使用电缆传输数据的总时延为 150＋5＝155(ms)；数据在卫星中的传输时延为 3Kb/60Kbps＝50ms，而其传播时延为 36 000km×2/(3×10^5 km/s)＝240ms(**注意**：卫星传输数据不是地面直接传输，而是要通过空中的卫星转发器转发，因此，卫星传输的距离近似为卫星距离地面高度的 2 倍)；因此使用卫星传输数据的总时延 50＋240＝290(ms)。因此本例使用电缆传输数据时间较短。

3.3 数据传输方式

在数据通信系统中，通信信道为数据的传输提供了各种不同的通路。对应于不同类型的信道，数据传输采用不同的方式，如并行传输和串行传输方式，单工、半双工和全双工通信方式，异步传输和同步传输方式，基带传输和频带传输方式等。

3.3.1 并行传输和串行传输

在计算机内部各部件之间，计算机与各种外部设备之间以及计算机与计算机(或终端)之间都是以数据传输的方式实现通信的。依据传输线数目的多少，可以将数据传输方式分为并行传输和串行传输，并行传输用于短距离、高速率的通信，串行传输用于长距离、低速率的通信。

1. 并行传输

在并行传输中，一般至少有 8 个数据位同时在两台设备之间进行传输，如图 3.4 所示。

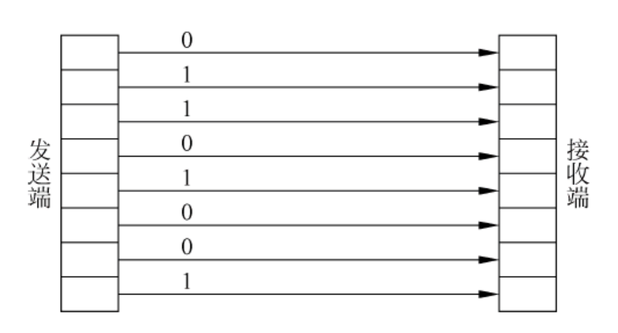


图 3.4　并行传输

并行传输是指数字信号以成组的方式在多个并行信道上进行的传输，数据由多条数据线同时传送与接收，每个比特使用单独的一条线路。

并行传输的优点在于传输速率快，发收双方间不存在字符同步的问题；缺点是需要多个并行信道，增加了设备的成本，而且并行线路的电平相互干扰也会影响传输质量，不适合做较长距离的通信。并行传输主要用于计算机内部或同一系统设备间的通信。常见的并行传输如计算机与打印机之间的数据传输。

2. 串行传输

并行传输需要 8 条以上的数据线，这对于近距离的数据传输来说，其费用还可以负担，但当进行远距离数据传输时，采用这种方式费用就太高了。所以，在数据通信系统中，较远距离的通信采用的是另一种传输方式——串行传输方式。

在串行传输中，发送端和接收端有一条数据线相连，各数据位依次串行通过该线路。如图 3.5 所示，源数据站向目的数据站发出 01101001 的串行比特流。发收两端一次只能发送或接收一个数据位，因此所需数据线数目大大减少，各数据位依次串行地通过通信线路。由于在计算机内部总线上传输的是并行数据，要与外部设备进行串行通信，在发送端就需要把并行数据转换成串行数据，在接收端还需将串行数据转换成并行数据，计算机内部的串行通信适配器负责进行串行数据和并行数据的转换。在计算机局域网中，计算机之间也是串行传输，网卡就负责串行数据和并行数据的转换工作。

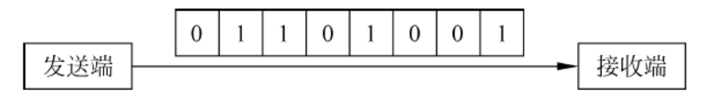


图 3.5　串行传输

相对于并行传输，串行传输的效率低，传输速率慢，但由于只有一条信道，减少了设备的成本，且易于实现和维护。串行传输适用于覆盖面很广的公共电话网络系统，所以在现行的计算机网络通信中，串行通信应用非常广泛。

3.3.2　单工、半双工和全双工通信

1. 单工通信

单工通信是指在两个通信设备间，信息只能沿着一个方向传输。

采用单工通信时，在通信设备双方中，一方为发送设备，另一方为接收设备，如图 3.6(a)所

示。广播和电视节目的传送以及寻呼系统都属于单工通信的例子。

2. 半双工通信

半双工通信是指两个通信设备间的信息交换可以双向进行,但不能同时进行。也就是说,在同一时刻仅能使信息在一个方向上传输,如图3.6(b)所示。

半双工通信设备的两端要求既要有发送设备,又要有接收设备,因此该方式需要具有信道转换能力,通常用软件控制换向,换向过程中存在换向的延迟时间问题,也可以采用人工操作机械开关的方法进行控制。典型的例子是对讲机或计算机与终端的通信。

3. 全双工通信

全双工通信是指两个通信设备间可以同时进行两个方向上的信息传输,如图3.6(c)所示。

通信双方应同时具有发送和接收的功能,与通信站相接的传输设备和传输控制协议必须提供全双工的工作方式,同时还应对缓存器作特殊的考虑。平时使用的手机类似于全双工通信。

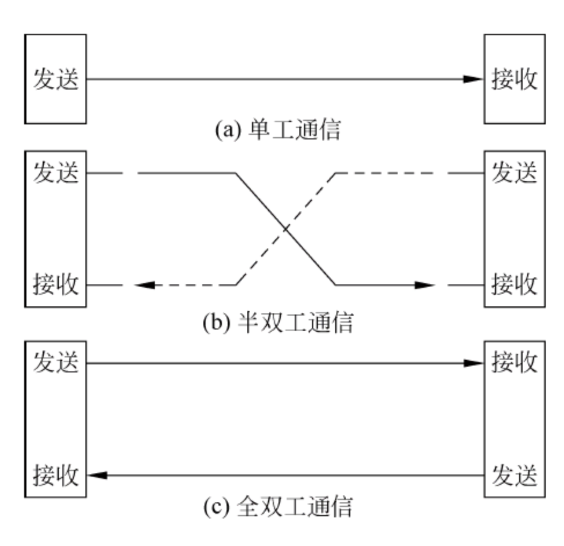


图3.6 三种通信方式

3.3.3 异步传输和同步传输

在数据通信中,有一个问题是必须解决的,那就是同步问题。同步问题就是发送方发出数据后,接收方如何从接收到的连续不断的信号中识别出数据的开始位置和结束位置。目前,在串行传输中所采用的同步方式有两种:一是异步传输方式;二是同步传输方式。

1. 异步传输方式

异步传输方式又称为起止式同步方式,它是以字符为单位进行传输的,即每个字符都独立传输。每个字符在传输时都在字符前加上起始位和在字符后加上结束位,以表示一个字符的开始和结束。

一般起始位信号的长度规定为1b(位)的宽度,极性为0,结束位信号可以为1b、1.5b或2b的宽度,极性为1,其长度的选取与所采用的传输代码类型有关。起始位和结束位的作用是实现字符同步,字符之间的间距是任意的,但发送一个字符时,每个字符包含的位数都是相

同的，且每一位占用的时间长度是双方约定好的，并且保持各位都恒定不变，如图 3.7 所示。

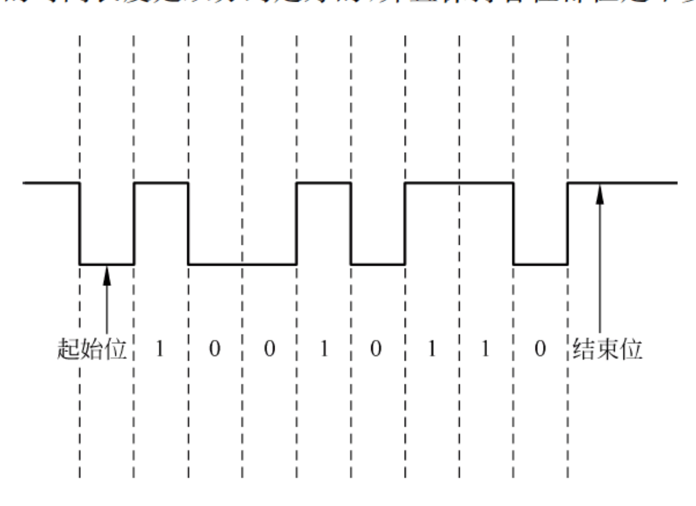


图 3.7 异步传输方式

在异步传输方式中字符可以被单独发送或连续发送，字符与字符的间隔期间可以连续发送 1 状态，当不传字符时，不要求收发时钟同步，仅在传输字符时，收发时钟才需要在字符的每一位上同步。同步的具体过程是：若发送端有信息要发送时，即将信号从不发送信息的 1 状态转到起始态 0，接收端检测出这种信号状态的改变时，就利用该信号的反转启动接收时钟，以实现收发时钟的同步。同理，接收端一旦收到结束位，就将定时器复位以准备接收下一个字符。

异步传输方式的优点是每一个字符本身就包括了本字符的同步信息，不需要在线路两端设置专门的同步设备；缺点是每发送一个字符就要添加一对起止信号，增加了线路开销，传输速率低。异步传输方式常用于小于或等于 1200bps 的低速率数据传输中，目前仍在广泛使用。

2. 同步传输方式

同步传输方式是以固定的时钟节拍来连续串行发送数字信号的一种方法。在数字信息流中，各位的宽度相同，且字符顺序相连，字符之间没有间隙。为使接收方能够从连续不断的数据流中正确区分出每一位(比特)，则需要先建立收发双方的同步时钟。实际上，在同步传输方式中，不管是否传送信息，要求收发两端的时钟都必须在每一位上保持一致。因此，同步传输方式又常被称为比特或位同步。

在同步传输中，数据的发送一般是以一组字符或比特流为单位进行的。为了使接收方容易确定数据组的开始和结束，需要在每组数据的前后加上特定字符作为起始标志和结束标志，同时还可以用这些标志来区分和隔离连续传输的数据。特定标志字符一般随不同的规程而有所不同。例如，在面向比特的高级数据链路控制规程 HDLC 中，采用比特串 01111110 作为起始标志和结束标志，如图 3.8 所示。在暂时没有信息传输时，连续发送 01111110 使接收端可以一直保持和发送端同步。

实现同步传输方式中的收发时钟同步的方法有两种：外同步法和自同步法。外同步法就是在传输线中增加一根时钟信号线以连接到接收设备的时钟上，在发送数据信号前，先向接收端发一串同步时钟脉冲，接收端则按照这个频率来调整自己的内部时钟，并把接收时钟重复频率锁定在同步频率上，该方法适用于近距离传输。自同步法是让接收方从接收的数

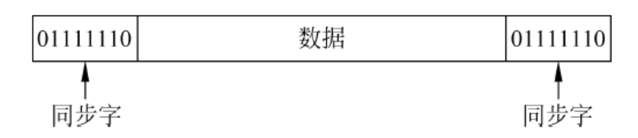


图 3.8 HDLC 中的同步传输方式

据流中直接提取同步信号，以获得与发送时钟完全相同的接收时钟，该方法常用于远距离传输。

同步传输克服了异步传输方式中的每一个字符都要附加起始信号和结束信号的缺点，具有较高的效率，但实现较为复杂，常用于大于 2400bps 速率的传输。

3.3.4 基带传输和数字信号编码

1. 基带传输

在数据通信中，由计算机或终端等数字设备产生的、未经调制的数字数据相对应的电脉冲信号通常呈矩形波形式，即表示计算机中二进制数据比特序列的数据信号是典型的矩形脉冲信号，这个矩形脉冲信号就是基带信号。基带信号所占有（固有）的频率范围称为基本频带，简称基带。在通信信道中直接传输这种基带信号的传输方式就是基带传输，它将占用线路的全部带宽，也称为数字基带传输。

2. 数字数据的数字信号编码

数字数据的数字信号编码问题就是要解决数字数据的数字信号表示问题，数字数据可以由多种不同形式的电脉冲信号的波形来表示，数字信号是离散的电压或电流的脉冲序列，每个脉冲代表一个信号单元（或称码元）。最普遍且最容易的方法是用两种码元分别表示二进制数字符号 0 和 1，每位二进制符号和一个码元相对应。表示二进制数字的码元的形式不同，产生的编码方法也不同，这里主要介绍曼彻斯特编码和差分曼彻斯特编码。

曼切斯特编码的编码方法是将每一个码元再分成两个相等的间隔，在时刻中间发生跳变。当为 0 时，在间隔的中间时刻，从低电平变为高电平；当为 1 时，在间隔的中间时刻，从高电平变为低电平。这种编码的特点就是在每一个码元时间间隔内，都有一次电平的跳转，对提取位同步信号非常有利。以太网中采用的就是这种编码技术。

差分曼切斯特编码的编码方法是在每一个码元时间间隔内，无论为 0 或为 1，在间隔的中间都有电平的跳转。但当为 0 时，间隔开始时刻有跳转；当为 1 时，间隔开始时刻无跳转。与曼切斯特编码的不同之处在于每位中间的跳转作为同步时钟信号，而取值是 0 还是 1 则根据每一位的起始处有没有变化来判断。令牌环网中采用的就是这种编码。曼切斯特编码和差分曼切斯特编码分别如图 3.9(a)和图 3.9(b)所示。

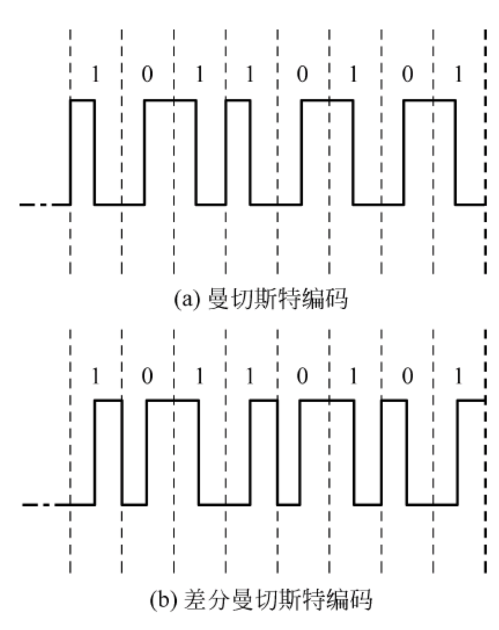


图 3.9 曼切斯特编码和差分曼切斯特编码

3.3.5 频带传输和模拟信号编码

1. 频带传输

由于基带信号频率很低，含有直流成分，在远距离传输过程中信号功率的衰减或干扰将造成信号减弱，使得接收方无法接收，因此基带传输不适合于远距离传输；另外，远距离通信信道多为模拟信道，所以，在远距离传输中不采用基带传输而是采用频带传输。频带传输就是先将基带信号(数字信号)进行调制后变换成便于在模拟信道中传输、具有较高频率范围的信号，这种信号称为频带信号(也叫模拟信号)，再将这种频带信号在信道中传输。在接收端再将该频带信号通过解调还原成基带信号。基带信号和频带信号的变换是由调制解调技术完成的，完成调制、解调的设备叫作调制解调器。计算机网络系统的远距离通信通常都是采用频带传输。

2. 数字数据的模拟信号编码

我们已经知道，在计算机网络的远程通信中通常采用频带传输。若要将基带信号进行远程传输，要先将其变换为频带信号，再在模拟信道上进行传输，这个变换就是数字数据的模拟信号的编码过程(即调制过程)。

所谓调制就是进行波形变换，利用基带信号对高频震荡载波的参量进行修改。最常用的载波是正弦波，假设振幅为1、频率为 f、初相位为 φ，则对应的数学表达式为 $u(t)=\sin(ft+\varphi)$。通过对载波的振幅、频率和初相位进行修改，分别对应了3种最基本的调制方法：调幅、调频和调相，如图3.10所示。

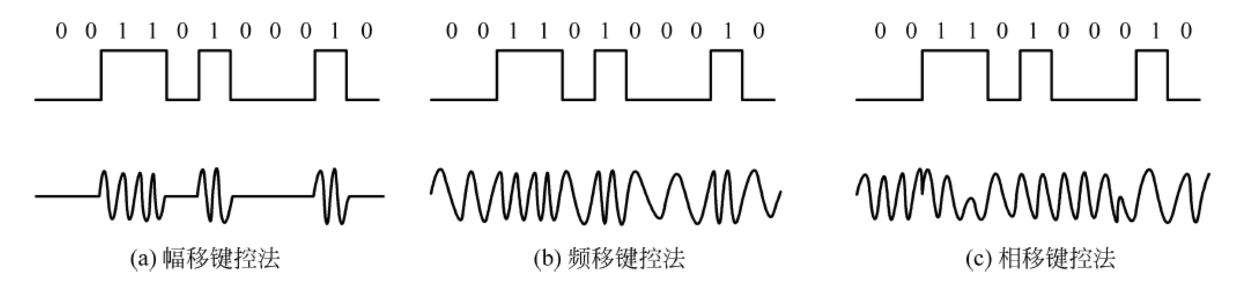


图3.10 数字数据的模拟信号调制

(1) 调幅(AM)。载波的振幅随基带信号的变化而变化，如0对应于无载波输出，即振幅为0；而1对应于有载波输出，即振幅为1，对应的数学表达式为

$$u(t)=\begin{cases}\sin(ft+\varphi)\\0\end{cases}$$

幅度调制又叫幅移键控(ASK)。在幅移键控方式下，用载波频率两个不同的振幅来表示两个二进制值。在有些情况下，用振幅恒定载波的存在来表示一个二进制数字，而用载波的不存在表示另一个二进制数字。ASK方式容易受增益变化的影响，因此，是一种效率相当低的调制技术。

(2) 调频(FM)。载波的频率随基带信号的变化而变化，如0对应于频率 f_1，而1对应于频率 f_2，对应的数学表达式为

$$u(t)=\begin{cases}\sin(f_1 t+\varphi)\\ \sin(f_2 t+\varphi)\end{cases}$$

频率调制也叫频移键控(FSK)。在频移键控方式下,用载波频率附近的两个不同频率来表示两个二进制值。这种方案比起 ASK 方式来,不容易受干扰的影响,一般也用于高频(3～30 MHz)的无线电传输,它甚至也能用于较高频率使用同轴电缆的局部网络。

(3) 调相(PM)。载波的初相位随基带信号的变化而变化,如 0 对应于相位 180°,而 1 对应于相位 0°,对应的数学表达式为

$$u(t)=\begin{cases}\sin(ft+0^\circ)\\ \sin(ft+180^\circ)\end{cases}$$

相位调制也叫相移键控(PSK)。在相移调制中,振幅和频率为常量,但通过控制或改变正弦载波信号的相位来表示二进制数据。根据使用相位的绝对值还是相位的相对偏移来表示二进制数据,我们将相位调制分为绝对调相和相对调相;按照对一个完整周期的相位等分方式,我们将相位调制分为二相制、四相制、八相制、十六相制……

3.3.6 模拟数据的数字信号编码

模拟数据的数字信号编码常用方法有脉冲编码调制(PCM)和增量调制(ΔM),现以 PCM 方法为例介绍。PCM 方法以取样定理为基础,将模拟数据数字化。例如,对音频信号进行数字化编码,一般包括取样、量化和编码 3 个过程。

(1) 取样。取样是指在每隔固定长度的时间点上抽取模拟数据的瞬时值,作为从这一次取样到下一次取样之间该模拟数据的代表值。根据取样定理,当取样的频率 F 大于或等于模拟数据的频带宽度(模拟信号的最高变化频率 F_{max})的 2 倍(即 $F \geqslant 2F_{max}$)时,所得的离散信号可以无失真地代表被取样的模拟数据。取样的结果是变连续的模拟信号为离散信号。取样也可以称为抽样或采样。

(2) 量化。量化就是把取样得到不同的离散幅值,按照一定量化级转换为对应的数据值,并取整数,得到离散信号的具体数值。所取的量化级越高,表示离散信号的精度越高。

(3) 编码。编码是将量化后的离散值转换为一定位数的二进制数值。通常,当量化级为 N 时,对应的二进制数为 $\log_2 N$。

3.4 多路复用技术

在远距离通信中,为了高效、合理地利用传输介质,通常采用多路复用技术,人们把利用一条物理信道同时传输多路信号的过程称为多路复用。多路复用技术是使多路数据信号共同使用一条线路进行传输的技术,使多台计算机或终端设备共享信道资源,提高信道的利用率。特别是在远距离传输时,可大大节省电缆成本、安装与维护费用。实现多路复用功能的设备是多路复用器。多路复用技术如图 3.11 所示。

多路复用技术通常有频分多路复用技术、时分多路复用技术、波分多路复用技术和码分

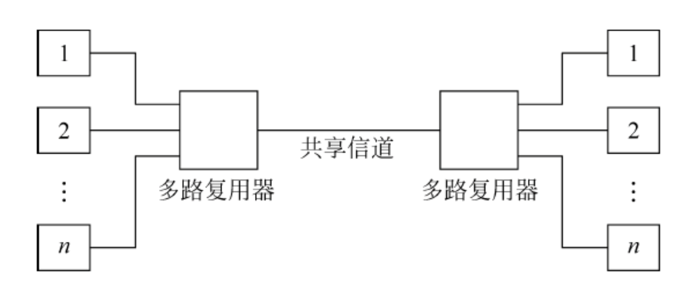


图 3.11　多路复用技术

多路复用技术。

3.4.1　频分多路复用

频分多路复用技术(FDM)是按照不同频率来区分信号的一种方法,将传输频带划分为若干个较窄的频带,每个频带传送一路信号,形成一个子信道。一个具有一定带宽的线路可以划分为若干个频率范围,相互之间没有重叠,同时,为了避免两个相邻频段的相互干扰,频段之间必须保留一定的缝隙,称为保护频带。这样,频分多路复用的所有用户在同样的时间内占用不同的频带资源。频分多路复用技术如图 3.12 所示。

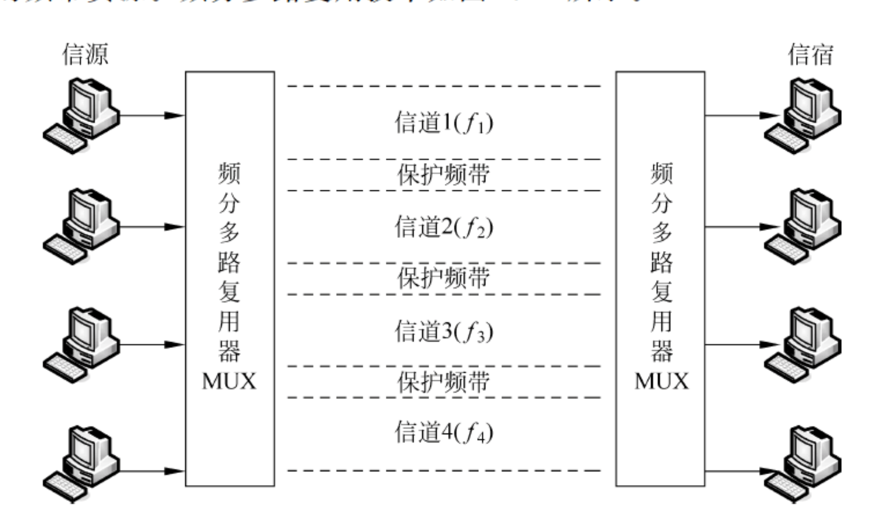


图 3.12　频分多路复用技术

频分多路复用常用于模拟信号的传输,如收音机、电视机等,也用于宽带网络。载波电话通信系统是频分多路复用的典型例子。

3.4.2　时分多路复用

时分多路复用技术(TDM)是将通信信道传输数据的时间划分成等长的时分复用帧(即 TDM 帧),每一个 TDM 帧再划分成若干等长的时间片,每一个时分多路复用的用户在每个 TDM 帧中占用固定序号的时间片来使用公共线路,在其占用的时间片内,信号独自使用信道的全部带宽。时分多路复用技术如图 3.13 所示。

从图 3.13 可以看出,一个用户所占用的时间片是周期出现的,这个周期就是一个时分

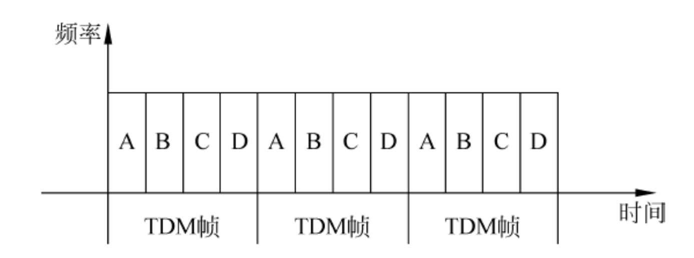


图 3.13 时分多路复用技术

多路复用帧的长度。时分多路复用技术的优点是技术比较成熟；缺点是不够灵活，如当用户在某一段时间暂时无数据传输时（如用户正在键盘上输入数据或正在浏览屏幕上的信息），也只能让已经分配到手的子信道空闲着，而其他用户却不能使用这个暂时空闲的信道资源。统计时分多路复用技术就是一种改进的时分多路复用技术，它能明显地提高信道的利用率。

统计时分多路复用（STDM）是使用 STDM 帧来传输数据的，但每一个 STDM 帧中划分的时间片的数目要小于进行复用的用户数，每一帧中的时间片不再是固定分配给某个用户，而是按需动态地给每个用户分配时间片。统计时分多路复用技术如图 3.14 所示。统计时分多路复用又称为异步时分多路复用，而普通的时分多路复用则称为同步时分多路复用。需说明的是，这里的帧与数据链路层的帧不是一个概念。

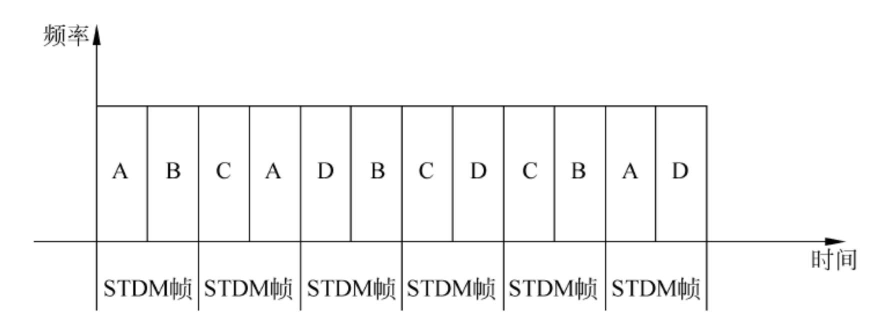


图 3.14 统计时分多路复用技术

时分多路复用技术通常用于数字信号的传输，也可用于模拟信号的传输。

时分多路复用在任一时刻只传送一种信号，多路信号分时地在信道中进行传输；而频分多路复用是在任一时刻，同时传送多路信号，各路信号占用的频带不同。

3.4.3 波分多路复用

波分多路复用技术（WDM）就是光的频分复用，是把光波波长分割复用，在一根光纤中同时传输多波长的光信号的一种技术。

波分多路复用技术的基本原理是在发送端将不同的光信号组合起来（即复用过程），然后耦合到光缆线路上，再用一根光纤进行传输；在接收端将组合波长的光信号区分开来（即解复用过程），再通过进一步处理恢复出原信号后送入不同终端。波分多路复用技术实质上是利用了光具有不同波长的特征。WDM 的原理十分类似于 FDM，不同的是，它利用波分多路复用设备将不同信道的信号调制成不同波长的光，并复用到光纤信道上。在接收方，采

用波分设备分离不同波长的光。相对于电多路复用器，WDM 发送端和接收端的器件分别称为分波器和合波器。

光波多路复用技术除 WDM 外，还有密集波分多路复用技术（DWDM），光纤的密集波分多路复用技术可极大地增加光纤信道的数量，从而充分利用光纤的潜在带宽，是计算机网络今后使用的重要技术。

3.4.4　码分多路复用

码分多路复用技术（CDM）是一种用于移动通信系统的新技术，笔记本电脑和掌上电脑等移动性计算机的联网通信将会大量使用码分多路复用技术。码分多路复用的基础是微波扩频通信，其特点是频率和时间资源均为共享。因此，在频率和时间资源紧缺的情况下，CDM 将独具魅力，越来越受到人们的关注。

3.5　传输介质

传输介质也称传输媒体，泛指计算机网络中用于连接各台计算机的物理媒体，特指用来连接各个通信处理设备的物理介质。传输介质是构成物理信道的重要组成部分，计算机网络中使用各种传输介质来组成物理信道。

3.5.1　传输介质的分类

传输介质包括有线传输介质和无线传输介质两大类。有线传输介质将信号约束在一个物理导体之内，如双绞线、同轴电缆和光纤等，故又被称为有线介质；而无线传输介质如无线电波、红外线、激光等不能将信号约束在某个空间范围之内，故被称为无线介质。究竟选择哪一种传输介质，必须考虑到价格、安装难易程度、容量、抗干扰能力、衰减等方面的因素，同时还要根据具体的运行环境全面考虑。

3.5.2　有线传输介质

1. 双绞线

双绞线（Twisted Pair，TP）是目前使用最广泛、价格最低廉的一种有线传输介质。Twisted 源于双绞线电缆的内部结构。双绞线在内部由若干对（通常是 1 对、2 对或 4 对）两两绞在一起的相互绝缘的铜导线组成，导线的典型直径为 1mm 左右（通常在 0.4～1.4mm）。之所以采用这种两两相绞的绞线技术，是为了抵消相邻线对之间所产生的电磁干扰以及减少线缆端接点处的近端串扰。

双绞线既可以传输模拟信号，也可以传输数字信号。用双绞线传输数字信号时，它的数据传输速率与电缆的长度有关。距离短时，数据传输速率可以高一些。典型的数据传输速率为 10Mbps、100Mbps 和 1000Mbps。

双绞线是把两根绝缘铜线拧成有规则的螺旋形。双绞线的抗干扰性较差,易受各种电信号的干扰,可靠性差。若把若干对双绞线集成一束,并用结实的保护外皮包住,就形成了典型的双绞线电缆。把多个线对扭在一起可以使各线对之间或其他电子噪声源的电磁干扰最小。

用于网络的双绞线和用于电话系统的双绞线是有差别的。

双绞线主要分为两类,即非屏蔽双绞线(Unshielded Twisted-Pair,UTP)和屏蔽双绞线(Shielded Twisted-Pair,STP)。与UTP相比,STP由于采用了良好的屏蔽层,所以抗干扰性较好,但由于价格较贵,因此在实际组网中用的不是很多。

到目前为止,EIA/TIA已颁布6类(Category,简写为Cat)线缆的标准。

(1) Cat 1类线:可用于电话传输,但不适合数据传输,这一级电缆没有固定的性能要求。

(2) Cat 2类线:可用于电话传输和最高为4Mbps的数据传输,包括4对双绞线。

(3) Cat 3类线:可用于最高为10Mbps的数据传输,包括4对双绞线,常用于10Base-T以太网。

(4) Cat 4类线:可用于16Mbps的令牌环网和大型10Base-T以太网,包括4对双绞线。其测试速率可达20Mbps。

(5) Cat 5类线:可用于100Mbps的快速以太网,包括4对双绞线。

(6) Cat 6类线:适用于1000Mbps的1000Base-TX,支持高达1000Mbps的数据通信。

双绞线使用RJ-45接头连接计算机的网卡或集线器等通信设备。组建局域网所用的双绞线是一种由4对线(即8根线)组成的,其中每根线的材质包括铜线和铜包的钢线两类。

一般来说,双绞线电缆中的8根线是成对使用的,而且每一对都相互绞合在一起,绞合的目的是为了减少对相邻线的电磁干扰。

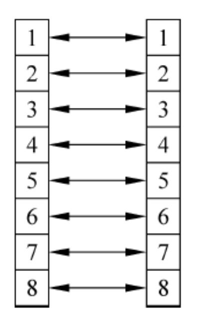
目前,在局域网中常用到的双绞线是非屏蔽双绞线(UTP),它又分为3类、4类、5类、超5类、6类和7类双绞线。

双绞线的这8根线的引脚定义如表3.1所示,Tx代表发送端,Rx代表接收端。

表3.1 双绞线各引脚定义

线号	1	2	3	4	5	6	7	8
线色	白橙	橙	白绿	蓝	白蓝	绿	白褐	褐
作用	Tx+	Tx−	Rx+			Rx−		

在局域网中,双绞线主要是用来连接计算机网卡到集线器或通过集线器之间接口的级联,有时也可直接用于两个网卡之间的连接或不通过集线器级联口之间的级联,但它们的接线方式各有不同,如图3.15所示。



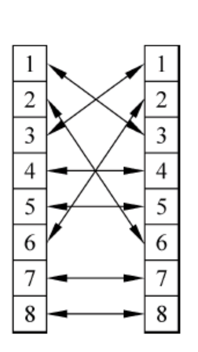


图3.15 双绞线的使用

双绞线具有以下特性。

(1) 物理特性:铜质线芯,传导性能良好。

(2) 传输特性:可用于传输模拟信号和数字信

号，对于模拟信号，5～6km需要一个放大器；对于数字信号，2～3km需要一个中继器。双绞线的带宽达268kHz。对于模拟信号，可用频分多路复用技术把它分成24路来传输音频模拟信号，根据目前的Modem技术，若使用相移键控法PSK，每路可达9600bps以上，这样，在一条24路的双绞线上，总传输速率可达230Kbps。对于数字信号，使用T1线路总传输速率可达1.544Mbps。达到更高传输速率也是可能的，但与距离有关。

(3) 对于局域网(10Base-T和100Base-T总线)，传输速率可达10～100Mbps。常用的3类双绞线和5类双绞线电缆均由4对双绞线组成，3类双绞线的传输速率可达10Mbps，5类双绞线的传输速率可达100Mbps，但与距离有关。

(4) 连通性：可用于点到点连接或多点连接。

(5) 地理范围：对于局域网，传输速率为100Kbps时，可传输1km；传输速率为10～100Mbps时，可传输100m。

(6) 抗干扰性：低频(10kHz以下)的抗干扰性能强于同轴电缆，高频(10～100kHz)的抗干扰性能弱于同轴电缆。

(7) 相对价格：比同轴电缆和光纤便宜得多。

2. 同轴电缆

同轴电缆是由一根空心的外圆柱形的导体围绕着单根内导体构成的。内导体为实芯或多芯硬质铜线电缆，外导体为硬金属或金属网。内外导体之间由绝缘材料隔离，外导体外还有外皮套或屏蔽物。

同轴电缆可以用于长距离的电话网络、有线电视信号的传输通道以及计算机局域网。50Ω的同轴电缆可用于数字信号发送，称为基带；75Ω的同轴电缆可用于频分多路复用转换的模拟信号发送，称为宽带。在抗干扰性方面，对于较高的频率，同轴电缆优于双绞线。

同轴电缆的中央是铜质的芯线(单股的实心线或多股的绞合线)，铜质的芯线外包一层绝缘层，绝缘层外是一层网状编织的金属丝作外导体屏蔽层(可以是单股的)，屏蔽层把电线很好地包起来，再往外就是外包皮的保护塑料外层了，如图3.16所示。

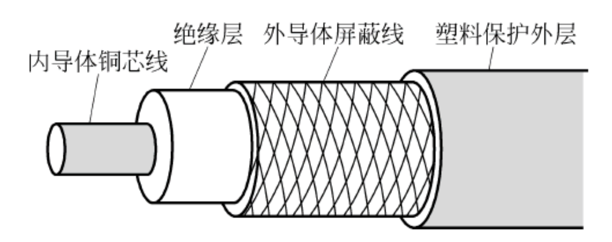


图3.16　同轴电缆的结构

目前，经常用于局域网的同轴电缆有两种：一种是专门用在符合IEEE 802.3标准以太网环境中阻抗为50Ω的电缆，只用于数字信号发送，称为基带同轴电缆；另一种是用于频分多路复用FDM的模拟信号发送，阻抗为75Ω的电缆，称为宽带同轴电缆。

同轴电缆具有以下特性。

(1) 物理特性：单根同轴电缆的直径为1.02～2.54cm，可在较宽频的范围内工作。

(2) 传输特性：基带同轴电缆仅用于数字传输，阻抗为50Ω，并使用曼彻斯特编码，数据传输速率最高可达10Mbps。宽带同轴电缆可用于模拟信号和数字信号传输，阻抗为75Ω，对于模拟信号，带宽可达300～450MHz。在CATV电缆上，每个电视通道分配6MHz带

宽，而广播通道的带宽要窄得多，因此，在同轴电缆上使用频分多路复用技术可以支持大量的视音频通道。

(3) 连通性：可用于点到点连接或多点连接。

(4) 地理范围：基带同轴电缆的最大距离限制在几千米，宽带同轴电缆的最大距离可以达几十千米。

(5) 抗干扰性：能力比双绞线强。

(6) 相对价格：比双绞线贵，比光纤便宜。

3. 光纤

光纤是一种细小、柔韧并能传输光信号的介质，一根光缆中包含有多条光纤。在光纤上用 1 表示有光脉冲信号，用 0 表示没有光脉冲信号。光纤通信系统是由光端机、光纤（光缆）和光纤中继器组成。光端机又分成光发送机和光接收机。而光纤中继器是用来延伸光纤或光缆的长度，防止光信号衰减。光发送机将电信号调制成光信号，利用光发送机内的光源将调制好的光波导入光纤，经光纤传送到光接收机。光接收机将光信号变换为电信号，经放大、均衡判决等处理后发送给接收方。

光纤和同轴电缆相似，只是没有网状屏蔽层。中心是光传播的玻璃芯，如图 3.17 所示。光纤分为单模光纤和多模光纤两类（所谓“模”，是指以一定的角度进入光纤的一束光）。

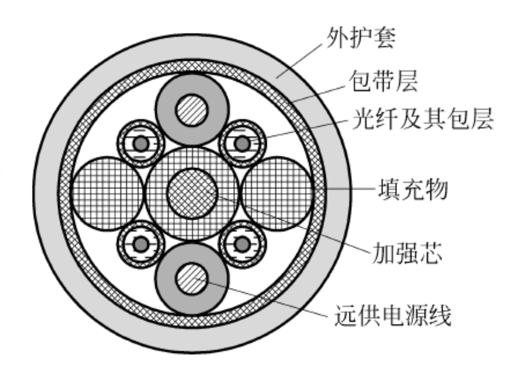


图 3.17　四芯光缆剖面示意图

正是由于光纤的数据传输速率高（目前已达到几 Gbps）、传输距离远（无中继传输距离达几十千米至上百千米）的特点，所以在计算机网络布线中得到了广泛的应用。目前，光缆主要是用于交换机之间、集线器之间的连接，但随着千兆位局域网应用的不断普及和光纤产品及其设备价格的不断下降，光纤连接到桌面也将成为网络发展的一个趋势。

光纤具有以下特性。

(1) 物理特性：在计算机网络中均采用两根光纤（一来一去）组成传输系统。按波长范围可分为 3 种：0.85μm 波长（0.8～0.9μm）、1.3μm 波长（1.25～1.35μm）和 1.55μm 波长区（1.53～1.58μm）。不同波长范围的光纤损耗特性也不同，其中 0.85μm 波长区为多模光纤通信方式，1.55μm 波长区为单模光纤通信方式，1.3μm 波长区有多模和单模两种光纤通信方式。

(2) 传输特性：光纤通过内部的全反射来传输一束经过编码的光信号，内部的全反射可以在任何折射指数高于包层媒体折射指数的透明媒体中进行。实际上光纤作为频率范围为 1014～1015Hz 的波导管，这一范围覆盖了可见光谱和部分红外光谱。光纤的数据传输速率可达 Gbps 级，传输距离达数十千米。目前，一条光纤线路上一般传输一个载波，随着技术的进一步发展，会出现实用的多路复用光纤。

(3) 连通性：采用点到点连接还有多点连接。

(4) 地理范围：在 6～8km 的距离内可以不用中继器传输，因此光纤适合于在几个建筑物之间通过点到点的链路连接局域网。

(5) 抗干扰性：不受噪声或电磁影响，适宜在长距离内保持高数据传输速率，而且能够

提供良好的安全性。

（6）相对价格：目前价格比同轴电缆和双绞线都贵。

3.5.3　无线传输介质

可以在自由空间利用电磁波发送和接收信号进行通信就是无线传输。地球上的大气层为大部分无线传输提供了物理通道，就是常说的无线传输介质。无线传输所使用的频段很广，人们现在已经利用了好几个波段进行通信，紫外线和更高的波段目前还不能用于通信。无线通信的方法有微波通信、激光通信和红外线通信。

1. 微波通信

微波通信系统有两种形式：地面系统和卫星系统。使用微波传输要经过有关管理部门的批准，而且使用的设备也需要有关部门允许才能使用。由于微波是在空间直线传播，如果在地面传播，地球表面是一个曲面，其传播距离受到限制，采用微波传输的站必须安装在视线内，传输的频率为4～6GHz和21～23GHz，传输距离一般只有50km左右。为了实现远距离通信，必须在一条无线通信信道的两个终端之间增加若干个中继站。中继站把前一站送来的信息经过放大后再送到下一站。通过这种“接力”通信，可以传输电话、电报、图像、数据等信息。采用卫星微波，卫星在发送站和接收站之间反射信号，传输的频率为11～14GHz。

目前，利用微波通信建立的计算机局域网也日益增多。由于微波是沿直线传输，所以长距离传输时要由多个微波中继站组成通信线路，而通信卫星可以看作是悬挂在太空中的微波中继站，可通过通信卫星实现远距离的信息传输。微波通信的主要特点是有很高的带宽（1～11GHz），容量大，通信双方不受环境位置的影响，并且不需要事先铺设电缆。

2. 激光通信

激光通信的优点是带宽更高、方向性好、保密性能好等，激光通信多用于短距离的传输。激光通信的缺点是其传输速率受天气影响较大。

3. 红外线通信

红外线通信不受电磁干扰和射频干扰的影响。红外无线传输建立在红外线光的基础上，采用光发射二极管、激光二极管或光电二极管来进行站点与站点之间的数据交换。红外无线传输既可以进行点到点通信，也可以进行广播式通信。但这种传输技术要求通信节点之间必须在直线视距之内，不能穿越墙。红外线传输技术数据传输速率相对较低，在面向一个方向通信时，数据传输速率为16Mbps。如果选择数据向各个方向上传输时，传输速率将不能超过1Mbps。

3.6　宽带接入技术

为了提高用户的上网速率，近年来已经有多种宽带接入技术进入用户的家庭。

3.6.1　拨号连接

拨号连接业务是用户通过拨打特服电话接入中国公众计算机网（CHINANET）的一种

低速上网方式。

采用拨号连接方式的用户需配备一台个人计算机，一套普通的拨号软件，一台调制解调器和一条电话线（普通电话或 ISDN 电话），通过拨打特别的接入号码进入互联网。拨号上网用户获得动态 IP 地址，普通电话拨号上网的最高传输速率可达 56Kbps，ISDN 电话拨号上网的最高传输速率可达 128Kbps。

使用拨号接入网络的客户端计算机要求必须安装有调制解调器（Modem），俗称“猫”。调制解调器分为内置和外置两种。使用拨号接入网络，用户上网速度很慢且通信链路十分不稳定。另外，拨号上网所需的电话线路因被调制解调器占用，所以无法提供电话语音服务。用户上网除了要支付网络流量所产生的费用外，还需要支付拨通接入号码所产生的电话费用。

现在，拨号连接方式基本已经很少被使用，只有在需要网络接入而其他接入方式又不能实现的情况下才会使用。

3.6.2 ADSL

ADSL 的全称是 Asymmetrical Digital Subscriber Loop，即非对称数字用户环路。ADSL 技术是运行在原有普通电话线上的一种新的高速宽带接入技术，它利用现有的一对电话铜线，为用户提供上下行非对称的传输速率（带宽）。非对称主要体现在上行传输速率（最高 640Kbps）和下行传输速率（最高 8Mbps）的非对称性上。上行（从用户到网络）为低速传输，可达 640Kbps；下行（从网络到用户）为高速传输，可达 8Mbps。最初，ADSL 主要是针对视频点播业务开发的，随着技术的发展，逐步成了一种较方便的宽带接入技术，是目前国内 ISP（因特网服务提供商）提供的主要接入服务的方式。

由于传统的电话线使用了 0～4kHz 的低频段进行语音传送，而电话线理论上有接近 2MHz 的带宽。ADSL Modem 采用频分多路复用（FDM）技术和回波消除（Echo Cancellation）技术在电话线上分隔有效带宽来实现多路信道。

频分多路复用技术在现有带宽中分配一段频带作为数据下行通道，同时分配另一段频带作为数据上行通道，下行通道通过时分多路复用（TDM）技术再分为多个高速信道和低速信道，同样在上行通道也由多路低速信道组成。

使用 ADSL 连接网络时，ADSL Modem 便在电话线上产生了 3 条信息通道：一个为标准电话服务的通道、一条速率为 640Kbps～1.0Mbps 的中速上行通道、一条速率为 1～8Mbps 的高速下行通道，并且这 3 条通道可以同时工作，而这一切都是在一根电话线上同时进行的。ADSL 使用了 26kHz 以后的高频带提供非常高的速率，它的具体工作流程是用户计算机产生的数字信号和电话产生的语音信号经滤波器编码后，信号通过电话线的中速上行通道传到电话局后再通过一个信号识别/分离器，如果是语音信号就传到电话交换机上；如果是数字信号就接入互联网上。互联网的返回数据在电信局端也同样与电话模拟信号进行混合，使用下行通道传输到用户端的滤波器时重新被分离为数字信号和模拟信号。

ADSL 接入技术具有以下特点。

（1）直接利用现有用户电话线，节省投资。

（2）享受超高速的网络服务，为用户提供上下行不对称的传输带宽。

(3) 节省费用,上网的同时可以打电话,互不影响,而且上网时不需要另交电话费。

(4) 安装简单,不需要另外申请增长率加线路,只需在普通电话线上加装 ADSL Modem,在计算机上装上网卡即可。

ADSL 接入 Internet 主要有虚拟拨号和专线接入两种方式。采用虚拟拨号方式的用户采用类似 Modem 和 ISDN 的拨号程序,在使用习惯上与拨号方式没什么不同。采用专线接入的用户只要开机即可接入 Internet。但是两种方式的网络结构是一样的,在客户端一般都需要有一台个人计算机、一台滤波器、一台 ADSL 调制解调器和一根电话线(目前来说调制解调器和滤波器一般都被整合为一台设备),如图 3.18 所示。

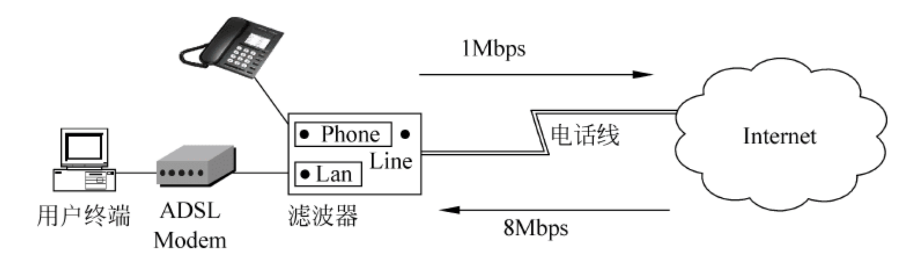


图 3.18　ADSL 接入方式示意图

1. ADSL 虚拟拨号接入

顾名思义,ADSL 虚拟拨号接入就是上网的操作,和普通 56Kbps Modem 拨号一样,有账号验证、IP 地址分配等过程。但 ADSL 连接的并不是具体的 ISP 接入号码(如 16300 或 16900),而是使用专门的 PPPOE 协议软件拨入 ADSL 虚拟专网接入的服务器。根据网络类型的不同又分为 ADSL 虚拟拨号接入和 Ethernet 局域网虚拟拨号方式两类;由于局域网虚拟拨号方式具有安装维护简单等特点,目前被广泛采用。

2. ADSL 专线接入

ADSL 专线接入是 ADSL 接入方式中的另一种,是采用一种直接使用 TCP/IP 协议类似于专线的接入方式。用户连接和配置好 ADSL Modem 后,在自己的个人计算机的网络管理中设置好相应的 TCP/IP 协议及网络参数(IP 和掩码、网关等都由局端事先分配好),计算机启动后用户端和局端会自动建立起一条链路。因此,ADSL 专线接入方式是以有固定 IP、自动连接等特点的类似专线的方式。具备固定的 IP 地址 ADSL 专线接入方式一般被 ISP 应用在需求较高的网吧、大中型企业宽带应用中,其费用相比虚拟拨号方式一般更高,个人用户一般很少考虑采用。

3.6.3　光纤同轴混合网 HFC

HFC 的全称是 Hybrid Fiber-Coaxial,即光纤和同轴电缆相结合的混合网络。HFC 接入方式是基于有线电视网络提供的。HFC 通常由光纤干线、同轴电缆支线和用户配线网络三部分组成,从有线电视台出来的信号先变成光信号在干线上传输;到用户区域后把光信号转换成电信号,经分配器分配后通过同轴电缆送到用户。

HFC 与早期 CATV 同轴电缆网络的不同之处,主要是在干线上用光纤传输光信号,在前端需完成电—光转换,进入用户区后要完成光—电转换。最初 HFC 网络是用来传输有线电视信号的,后来通过对现有有线电视网进行双向化改造。HFC 网络除了可以提供有线

电视节目外还可以提供电话、Internet 接入、高速数据传输和多媒体等业务，如图 3.19 所示。

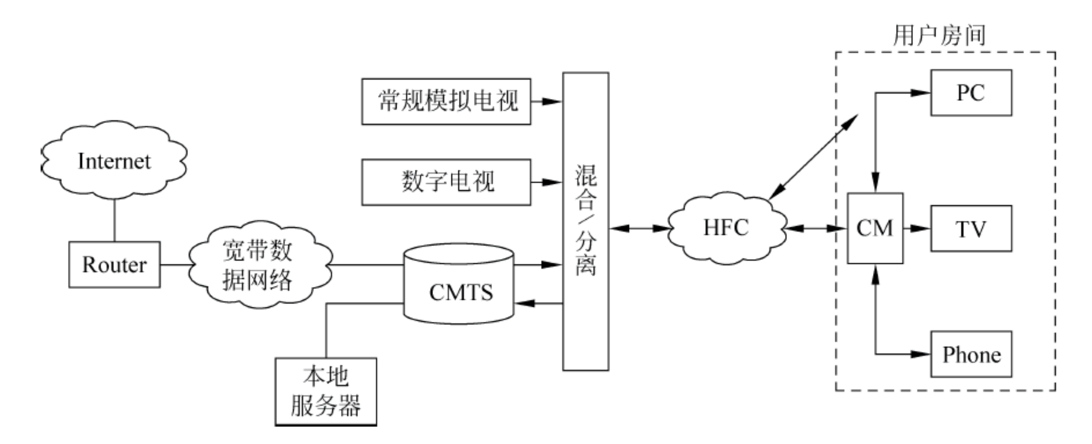


图 3.19 HFC 网络结构图

HFC 网络接入主要采用局端系统(Cable Modem Termination Sys)，完成数据到射频 RF 转换，并与有线电视的视频信号混合，送入 HFC 网络中。除了与高速网络连接外，也可以作为业务接入设备，通过 Ethernet 网口挂接本地服务器提供本地业务。

用户在接入 HFC 网络时，需要一台 Cable Modem 以及用户计算机。电缆调制解调器(Cable Modem，CM)是一种将数据终端设备(计算机)连接到 HFC，以使用户能进行数据通信，访问 Internet 等信息资源的设备。主要用于有线电视网进行数据传输。

Cable Modem 有两个接口，一个接室内墙上的有线电视端口；另一个与计算机相连。

Cable Modem 包含以下功能模块。

(1) 调制解调部分。

(2) 电视接收调谐、加密解密和协议适配等。

(3) 提供标准的以太网接口。

一个 Cable Modem 要在两个不同的方向上接收和发送数据，把上下行数字信号用不同的调制方式调制在双向传输的某一个 6MHz(或 8MHz)带宽的电视频道上。它把上行的数字信号转换成模拟射频信号，在有线电视网上传送。接收下行信号时，Cable Modem 把它转换为数字信号，以便计算机处理。使用 HFC 进行网络接入不需要使用任何拨号软件进行拨号，在用户端只要将线路连接正确，开机就可以接入互联网。

使用 HFC 接入方式只能在广播电视局进行申请，HFC 网络具有以下特点。

(1) 传输容量大，易实现双向传输。从理论上讲，一对光纤可同时传送 150 万路电话或 2000 套电视节目。

(2) 频率特性好，在有线电视传输带宽内无须均衡。

(3) 传输损耗小，可延长有线电视的传输距离，25km 内无须中继放大；光纤间不会有串音现象，不怕电磁干扰，能确保信号的传输质量。

但是，HFC 是在单向的基础上进行双向的改造来进行传输。由于它共享一条信道，它的带宽在用户量增加时，会不断减少，相互的干扰过大。没有一个网络在 Cable 上的用户超过 5000 个。目前，有线电视网在带宽共享方式、网络安全、网络管理等方面依然存在缺陷。而且从网络结构上来看，整个 HFC 用户网络都是属于同一个广播域，随着网络用户的增加

网络性能会迅速下降。

3.6.4 代理服务器接入

当从 ISP 接入用户的接入线路只有一条，而用户端有多台计算机需要对互联网进行访问时，使用代理服务器是最常用的手段。

1. 代理服务器基本概念

代理服务器英文全称是 Proxy Server，其功能就是代理网络用户去取得网络信息。形象地说，它是网络信息的中转站。

代理服务器是一台配备了两块以太网网卡的服务器。该服务器有一块以太网网卡接入 Internet，接入方式可以为城域网的 10/100M 以太网接口、ISDN、PSTN，或者是 ADSL；另一块以太网网卡一般和内部局域网互联，使用代理软件来进行代理业务处理。

连接 ISP 网络的以太网网卡要设置 ISP 提供的公共网络 IP 地址，用来在互联网上进行路由。而连接内部局域网的网卡要设置私有 IP 地址，这些私有 IP 地址主要是实现在内部网络的通信，不能在互联网上进行路由。

在局域网中的计算机需要访问外部网络时，该计算机的访问请求被代理服务器截获，代理服务器通过查找本地的缓存，如果请求的数据(如 WWW 页面)在缓存中可以查找到，则把该数据直接传给局域网络中发出请求的计算机；否则代理服务器访问外部网络，获得相应的数据，并把这些数据存入缓存，同时把该数据发送给发出请求的计算机。代理服务器缓存中的数据会随着内部网络计算机对互联网的访问而不断更新。一般在代理服务器上安装运行代理软件来实现内部网络的计算机对外部网络访问时的处理过程，常用的代理服务器软件有 SyGate、WinGate、CCProxy 等。

代理服务器的主要功能如下。

(1) 突破自身 IP 访问限制，访问国外站点。教育网、169 网等网络用户可以通过代理访问国外网站。

(2) 访问一些单位或团体内部资源，如某大学 FTP(前提是该代理地址在该资源的允许访问范围之内)，使用教育网内地址段免费代理服务器，就可以用于对教育网开放的各类 FTP 资源进行下载、上传以及查询共享等服务。

(3) 突破 ISP 的 IP 封锁：有很多被限制访问的网站是人为的，不同服务器对地址的封锁是不同的。使用相应的代理服务器可以实现合法访问。

(4) 提高访问速度：通常代理服务器都设置一个较大的硬盘缓冲区，当有外界的信息通过时，同时也将其保存到缓冲区中，当其他用户再访问相同的信息时，则直接由缓冲区中取出信息传给用户，以提高访问速度。

(5) 隐藏真实 IP：上网者也可以通过这种方法隐藏自己的 IP，免受攻击。

2. 代理服务器类型

代理服务器的类型很多，如 HTTP 代理、FTP 代理、SOCKS 代理等，每类代理各自都有其对应的功能。

(1) HTTP 代理：代理客户机的 HTTP 访问，主要是代理浏览器访问网页，它的端口一般为 80、8080、3128 等。

（2）FTP 代理：代理客户机上的 FTP 软件访问 FTP 服务器，它的端口一般为 21、212。

（3）RTSP 代理：代理客户机上的 Real Player 软件访问 Real 流媒体服务器的代理，其端口一般为 554。

（4）POP3 代理：代理客户机上的邮件软件用 POP3 方式收发邮件，其端口一般为 110。

（5）SOCKS 代理：SOCKS 代理与其他类型的代理不同，它只是简单地传递数据包，而并不关心是何种应用协议，所以 SOCKS 代理服务器比其他类型的代理服务器速度要快得多。SOCKS 代理又分为 SOCKS4 和 SOCKS5，二者不同的是 SOCKS4 代理只支持 TCP 协议（即传输控制协议），而 SOCKS5 代理则既支持 TCP 协议，又支持 UDP 协议（即用户数据包协议）；此外，还支持各种身份验证机制、服务器端域名解析等。例如，常用的聊天工具 QQ 在使用代理时就要求用 SOCKS5 代理，因为它需要使用 UDP 协议来传输数据。

以上是针对协议类型来对代理服务器进行的分类，实际使用当中代理软件其实包括了所有的提到的协议类型。一般常用的代理软件有 SyGate、WinGate、CCProxy 等。

课后习题

1. 术语解释

数据传输速率　信道带宽　频带传输　频分多路复用　时分多路复用　全双工通信　调制与解调　DTE　DCE

2. 信息是(　　)。

A. 消息　　　　B. 数据

C. 可以辨别的符号　　　　D. 经过加工处理的数据

3. 目前，在计算机网络系统中主要采用的复用方式是(　　)。

A. 波分多路复用　　　　B. 时分多路复用

C. 码分多路复用　　　　D. 频分多路复用

4. ADSL 的“非对称”性是指________________。

5. HFC 中的上行信号是指________________，下行信号是指________________。

6. 物理层要解决哪些问题？物理层的主要特点是什么？

7. 常用的传输媒体有哪几种？各有何特点？

8. 什么是曼彻斯特编码和差分曼彻斯特编码？其特点如何？

9. 什么是基带传输？什么是频带传输？两者在数据通信系统的组成上有什么区别？分别要解决什么样的关键问题？

10. 什么是数据通信？衡量数据通信质量的主要性能指标有哪些？

11. 有 600 MB(兆字节)的数据，需要从南京传送到北京。一种方法是将数据写到磁盘上，然后托人乘火车将这些磁盘捎去。另一种方法是用计算机通过长途电话线路（设信息传输速率是 2.4Kbps）传送此数据。试比较这两种方法的优劣。若信息传输速率为 33.6Kbps，其结果又如何？

12. 收发两端之间的传输距离为 1000km，信号在媒体上的传输速率为 2×10^8 m/s。试计算以下两种情况的发送时延和传播时延：

(1) 数据长度为 10^7 b,数据传输速率为 100Kbps;

(2) 数据长度为 10^3 b,数据传输速率为 1Gbps。

从以上计算结果可得出什么理论?

13. 假设信号在媒体上的传输速率为 2.3×10^8 m/s。媒体长度 l 分别为

(1) 10cm (网络接口卡);

(2) 100m (局域网);

(3) 100km (城域网);

(4) 5000km (广域网)。

试计算当数据率为 1Mbps 和 10Gbps 时,在以上媒体中正在传播的比特数。

14. 有一个点对点链路,长度为 50km。若数据在此链路上的传输速率为 2×108m/s,试问链路的宽带应为多少才能使传播时延和发送时延 100B 的分组的发送时延一样大?如果发送的时延是 512B 长的分组,结果又如何?

15. 条件如上题,但数据的传输速率改为 1Mbps。和上题的结果相比较,可以得出什么结论?

第4章 数据链路层

学习目的

在讨论了物理层协议与标准后，本章进一步讨论数据链路层协议与标准。本章将研究数据链路层的基本概念和服务功能，分析差错产生的原因与差错控制方法，讨论流量控制的方法以及典型的数据链路层协议。

学习要求

掌握：数据链路层的基本概念、功能与作用。

了解：数据传输过程中差错产生的原因与性质。

掌握：差错控制的作用和原理。

掌握：面向比特型数据链路层协议实例——HDLC。

掌握：Internet 中的数据链路层协议。

4.1 数据链路层概述

4.1.1 数据链路层的必要性

“链路”和“数据链路”是两个不同的概念。“链路”是指一条无源的点到点物理线路，中间没有任何其他的交换节点。网络上两台计算机在通信时，所传送的数据往往要经过许多中间交换节点，因而一条从源节点到目标节点的通路往往是由许多链路串接而成的，也就是说，一条链路往往只是一条通路的一部分。“数据链路”则是一条物理通路加上必要的数据传输规程或协议后所形成的逻辑连接。这种规程或协议控制着数据在链路上的传输。

至少有两个理由可用来说明数据链路层存在的必要性。首先是数据传输过程中的损坏与丢失问题。尽管物理层采取了一些必要的措施来减少信号传输过程中的噪声，但是数据在物理传输过程中仍然可能被损坏或丢失。由于物理层只关心原始比特流的传送，不考虑也不可能考虑所传输信号的意义和信息的结构，所以物理层不可能识别或判断数据在传输过程中是否出现了损坏或丢失，从而也谈不上采取相应的机制或方法进行补救。其次是收发双方的接收和发送速率不匹配引发的数据丢失问题。当数据发送方的发送能力大于数据接收方的接收能力时，接收方会因为来不及处理而产生数据溢出并导致数据丢失。然而，物理层并不考虑当发送站点发送速率过快而接收站点接收速率过慢时，应采取何种策略来控制发送站点的发送速率，以避免接收站点来不及处理而丢失数据。可见只有物理层的功能是不够的，位于物理层之上的数据链路层就是为了克服物理层的这些不足而建立的。

数据链路层旨在实现网络上两个相邻节点之间的无差错传输。它利用了物理层提供的

原始比特流传输服务，检测并校正物理层的传输差错，控制数据的传输流量，使在相邻节点之间构成一条无差错的链路，从而向网络层提供可靠的数据传输服务。

4.1.2　数据链路层的功能

数据链路层最基本的服务是将源机网络层来的数据可靠地传输到相邻节点的目标机网络层。为达到这一目的，数据链路层必须具备一系列相应的功能，主要包括：如何将数据组合成数据块，在数据链路层中将这种数据块称为帧，帧是数据链路层的传送单位；如何控制帧在物理信道上的传输，包括如何处理传输差错，如何调节发送速率以使之与接收方相匹配；在两个网络实体之间提供数据链路通路的建立、维持和释放管理。

4.1.3　数据链路层所提供的基本服务

通常，数据链路层有3种基本服务可供选择，即无确认的无连接服务（Unacknowledged Connectionless Service）、有确认的无连接服务（Acknowledged Connectionless Service）、有确认的面向连接服务（Acknowledged Connection-oriented Service）。

（1）在无确认的无连接服务方式下，两台相邻机器之间在发送数据帧之前，事先不建立连接，事后也不存在释放连接。源机器向目标机器发送独立的数据帧，而目标机器不对收到的帧作确认。对于线路上的噪声而造成的帧丢失，数据链路层将不作努力去恢复，而是将该工作留给上层（通常为传输层）去完成。这类服务通常适用于误码率很低的信道，如大多数局域网使用这种无确认的无连接服务方式。

（2）在有确认的无连接服务方式下，仍然不需要建立连接，源机器向目标机器发送独立的数据帧，但是接收站点要对收到的每一帧作确认，即在收到数据帧之后回送一个确认帧，而发送站点在收到确认帧之后才会发送下一帧。当在一个确定的时间段内没有收到确认帧时，发送方就认为所发送的数据帧丢失并自动重发此帧。自动重发可能会产生接收站点收到重复的数据帧的问题。有确认的无连接服务方式适用于像无线网之类的不可靠信道。

（3）在有确认的面向连接服务方式下，发送数据之前，需要首先建立连接，然后才会启动帧的传输。在发送数据阶段，为所传输的每一帧都要编上号，数据链路层提供相应的确认和流量控制机制来保证每一帧都只被正确接收一次，并保证所有帧都按正确的顺序被接收。当数据传输完成之后，还需要拆除或释放所建立的连接。也就是说，面向连接服务方式分3个阶段：链路建立阶段、数据传输阶段和链路拆除阶段。可以这么说，只有有确认的面向连接服务方式才真正为网络层提供了可靠的无差错传输服务。这类服务实现复杂度及代价很高，通常被用于误码率较高的不可靠信道，如某些广域网链路。

4.2　帧与成帧

为了实现上述数据链路层一系列功能，数据链路层必须使自己所看到的数据是有意义的，其中除了要传送的用户数据外，还要提供关于寻址、差错控制和流量控制所必需的信息，

而不再是物理层所谓的原始比特流。为此,数据链路层采用了被称为帧(Frame)的协议数据单元作为数据链路层的数据传送逻辑单元。不同的数据链路层协议的核心任务就是根据它所要实现的数据链路层功能来规定帧的格式。

4.2.1 帧的基本格式

尽管不同数据链路层协议给出的帧格式都存在一定差异,但它们的基本格式基本相同。图4.1给出了帧的基本格式,组成帧的那些具有特定意义的部分被称为域或字段(Field)。

帧开始	地址	长度/类型/控制	数据	FCS	帧结束

图4.1 帧的基本格式

(1)帧开始:用以指示一个帧的开始。

(2)地址:用于设备或机器的物理寻址,使得能够在多个相邻节点之间确定一个接收目标。

(3)长度/类型/控制:这个字段在不同的数据链路层协议中可以有不同的规定。或给出帧的长度信息,或给出帧的类型信息,或表明该帧为控制帧。长度通常以字节为单位;帧的类型主要包括提供数据传输的数据帧和提供链路控制与传输管理功能的控制帧。

(4)数据:承载的来自高层(网络层)的数据分组(Packet)。

(5)FCS:表示帧检验序列(Frame Check Sequence),该字段提供与差错检测有关的信息。

(6)帧结束:用以指示一个帧的结束。该字段与帧开始字段一起提供了数据流的定界,使得接收方可以正确识别数据流的开始与结束。

通常,“数据”字段之前的那些字段被统称为帧头(Head)部分,而“数据”字段之后的所有字段被称为帧尾(Trailer)部分。

4.2.2 成帧与拆帧

引入帧机制不仅可以实现相邻节点之间的可靠传输,还有助于提高数据传输的效率。例如,若发现接收到的某一个(或某几个)比特出错时,可以只对相应的帧进行特殊处理(如请求重发等),而不需要对其他未出错的帧进行这种处理;如果发现某一帧被丢失,也只要请求发送方重传所丢失的帧,从而大大提高了数据处理和传输的效率。但是,引入帧机制后,发送方的数据链路层必须提供将从网络层接收的分组(Packet)封装成帧的功能,即为来自上层的分组加上必要的帧头和帧尾部分,通常称此为成帧(Framing);而接收方数据链路层则必须提供将帧重新拆装成分组的拆帧功能,即去掉发送端数据链路层所加的帧头和帧尾部分,从中分离出网络层所需的分组。在成帧过程中,如果上层的分组大小超出下层帧的大小限制,则上层的分组还要被划分成若干个帧才能被传输。

发送端和接收端数据链路层所发生的帧发送和接收过程大致如下:发送端的数据链路层接收到网络层的发送请求之后,便从网络层与数据链路层之间的接口处取下待发送的分组,并封装成帧,然后经过下层物理层送入传输信道;这样不断地将帧送入传输信道就形成

了连续的比特流；接收端的数据链路层从来自其物理层的比特流中识别出一个一个的独立帧，然后利用帧中的 FCS 字段对每一个帧进行校验，判断是否有错误。如果有错误，就采取收发双方约定的差错控制方法进行处理；如果没有错误，就对帧实施拆封，并将其中的数据部分通过数据链路层与网络层之间的接口上交给网络层，从而完成相邻节点的数据链路层关于该帧的传输任务。

4.2.3 帧定界

为了使传输中发生差错后只将出错的有限数据进行重发，数据链路层将比特流组织成以帧为单位传送。帧的组织结构必须设计成使接收方能够明确地从物理层收到的比特流中对其进行识别，也即能从比特流中区分出帧的起始与终止，这就是帧定界要解决的问题，也称为帧同步。由于帧同步网络传输中很难保证计时的正确和一致，所以不能采用依靠时间间隔关系来确定一帧的起始与终止的方法。有 4 种常见的定界方法，即字节计数法、使用字符填充的首尾定界符法、用比特填充的首尾定界符法和违法编码法。

1. 字节计数法

这种帧同步方法以一个特殊字符表征一帧的起始，并以一个专门字段来标明帧的字节数。接收方可以通过对该特殊字符的识别从比特流中区分帧的起始，并从专门字段中获知该帧随后跟随的数据字节数，从而可确定帧的终止位置。

面向字节计数的同步规程的典型实例是 DEC 公司的数字数据通信报协议 DDCMP (Digital Data Communications Message Protocol)。DDCMP 采用的帧格式如图 4.2 所示。

8	14	2	8	8	8	16	16～131 064	16(b)
SOH	Count	Flag	Ack	Seg	Addr	CRC_1	Data	CRC_2

图 4.2 DDCMP 帧格式

格式中控制字符 SOH 标志数据帧的起始。Count 字段共有 14B，用以指示帧中数据段数据的字节数，数据段最大长度为 $8\times(2^{14}-1)=131\ 064$B，长度必须为字节(即 8b)的整倍数，DDCMP 协议就是靠这个字节计数来确定帧的终止位置。CRC_1、CRC_2 分别对标题部分和数据部分进行双重校验，强调标题部分单独校验的原因是：一旦标题部分中的 Count 字段出错，即丢失了帧边界划分的依据，将造成灾难性的后果。

2. 使用字符填充的首尾定界符法

该法用一些特定的字符来定界一帧的起始与终止。为了使数据信息位中出现的与特定字符相同的字符不被误判为帧的首尾定界符，可以在这种数据字符前填充转义控制字符以示区别，从而达到数据的透明性。带字符填充的首尾定界符法是在每一帧的开头用 ASCII 字符 DLE STX，在帧末尾用 ASCII 字符 DLE ETX。但是，如果在帧的数据部分也出现了 DLE STX 或 DLE ETX，那么接收端就会错误判断帧边界。为了不影响接收方对帧边界的正确判断，采用了填充字符 DLE 的方法。即如果发送方在帧的数据部分遇到 DLE，就在其前面再插入一个 DLE，这样数据部分的 DLE 就会成对出现。在接收方，若遇到两个连续的 DLE，则认为是数据部分，并删除一个 DLE。

3. 用比特填充的首尾定界符法

该法以一组特定的比特模式(如 01111110)来标志一帧的起始与终止。为了不使信息

位中出现的与该特定模式相似的比特串误判为帧的首尾标志，可以采用比特填充的方法。比如，采用特定模式 01111110，则对信息位中的任何连续出现的 5 个 1，发送方自动在其后插入一个 0，而接收方则做该过程的逆操作，即每收到连续 5 个 1，则自动删去其后所跟的 0，以此恢复原始信息，实现数据传输的透明性。比特填充很容易由硬件来实现，性能优于字符填充方法。

4. 违法编码法

该法在物理层采用特定的比特编码方法。例如，曼彻斯特编码方法，是将数据比特 1 编码成“高—低”电平对，将数据比特 0 编码成“低—高”电平对。而“高—高”电平对和“低—低”电平对在数据比特中是违法的。可以借用这些违法编码序列来定界帧的起始与终止。局域网 IEEE 802 标准中就采用了这种方法。违法编码法不需要任何填充技术，便能实现数据的透明性，但它只适用采用冗余编码的特殊编码环境。

由于字节计数法中 Count 字段的脆弱性（其值若有差错将导致灾难性后果）以及字符填充实现上的复杂性和不兼容性，目前较普遍使用是比特填充法和违法编码法。

4.3 差错控制

所谓差错，是指接收端收到的数据与发送端发出的数据出现不一致的现象。产生差错主要是因为在通信线路上噪声干扰的结果。根据噪声类型不同，可将差错分为随机错和突发错。热噪声所产生的差错称为随机错，冲击噪声（如电磁干扰、无线电干扰等）所产生的错误称为突发错。

差错的严重程度由误码率来衡量，误码率 P_e 等于错误接收的码元数与所接收的码元总数之比。显然，误码率越低，信道的传输质量越高，但是由于信道中的噪声是客观存在的，所以不管信道质量多高，都要进行差错控制。

4.3.1 差错控制的作用与机制

为了提高传输的准确性，采用了专门的校验错误方法，用来发现所产生的错误，并给出出现错误的信号或者校正错误。差错控制是采用可靠、有效的编码以减少或消除计算机通信系统中传输差错的方法，其目的在于提高传输质量。

为了有效地提高传输质量，一种方法是改善通信系统的物理性能，使误码的概率降低到满足要求的程度，但这种方法受经济和技术上的限制；另一种方法是差错控制，它是利用编码的手段将传输中产生的错码检测出来，并加以纠正。差错控制是数据通信中常用的方法。差错控制的主要作用是通过发现数据传输中的错误，采取相应的措施减少数据传输错误。差错控制的核心是对传输的数据信息加上与其满足一定关系的冗余码，形成一个加强的、符合一定规律的发送序列，所加入的冗余码称为校验码（Frame Check Sequence，FCS）。

校验码按功能不同被分为纠错码和检错码。纠错码不仅能发现传输中的错误，还能利用纠错码中的信息自动纠正错误，其对应的差错控制措施为自动前向纠错。汉明码（Hamming Code）为典型的纠错码，具有很高的纠错能力。检错码只能用来发现传输中的

错误，但不能自动纠正所发现的错误，需要通过反馈重发来纠错。常见的检错码有奇/偶校验码和循环冗余校验码。目前，计算机网络通信中大多采用检错码方案。

4.3.2　奇/偶校验码

奇/偶校验的规则是在原数据位后附加一个校验位，将值置为0或1，使附加该位后的整个数据码中1的个数成为奇数或偶数。使用奇数个1进行校验的方案被称为奇校验；对应于偶数个1的校验方案被称为偶校验。奇/偶校验有3种使用方式，即水平奇/偶校验、垂直奇/偶校验和水平垂直奇/偶校验。下面以奇校验为例进行介绍。

水平奇校验码是指在面向字符的数据传输中，在每个字符的7位信息码后附加一个校验位0或1，使整个字符中二进制位1的个数为奇数。例如，设传输字符的比特序列为1100001，则采用奇校验码后的比特序列形式为11000010。接收方在收到所传输的比特序列后，通过检查序列中的1的个数是否仍为奇数来判断传输是否发生了错误。若比特在传输过程中发生错误，就可能会出现1的个数不为奇数的情况。水平奇校验只能发现字符传输中的奇数位错，而不能发现偶数位错。例如，上述发送序列11000010，若接收端收到11001010，则可以校验出错误，因为有一位0变成了1；但是若收到11011010，则不能识别出错误，因为有两位0变成了1。不难理解，水平偶校验也存在同样的问题。

为了提高奇/偶校验码的检错能力，引入了水平垂直奇/偶校验，即由水平奇/偶校验和垂直奇/偶校验综合构成。

垂直奇/偶校验也称为组校验，是将所发送的若干个字符组成字符组或字符块，形式上相当于是一个矩阵，如图4.3所示，每行为一个字符，每列为所有字符对应的相同位。在这一组字符的末尾即最后一行附加一个校验字符，该校验字符中的第i位分别对应组中所有字符第i位的校验位。显然，如果单独采用垂直奇/偶校验，则只能检出字符块中某一列中的一位或奇数位错。

但是，如果同时采用了水平奇/偶校验和垂直奇/偶校验，既对每个字符作水平校验，同时也对整个字符块作垂直校验，则奇/偶校验码的检错能力可以明显提高。这种方式的奇/偶校验被称为水平垂直奇/偶校验，图4.4给出了一个水平垂直奇/偶校验的例子。但是从总体上讲，奇/偶校验方法的检错能力仍较差，虽然其实现方法简单。故这种校验一般只用于通信质量要求较低的环境。

字母	前7行为对应字母的ASCⅡ码，最后一行是垂直奇校验码(粗体)
a	1 1 0 0 0 0 1
B	1 1 0 0 0 1 0
c	1 1 0 0 0 1 1
d	1 1 0 0 1 0 0
e	1 1 0 0 1 0 1
f	1 1 0 0 1 1 0
g	1 1 0 0 1 1 1
校验位	**0 0 1 1 1 1 1**

图4.3　垂直奇/偶校验

字母	最后一行是垂直奇校验码，最后一列是水平奇校验编码(粗体)
a	1 1 0 0 0 0 1 **0**
B	1 1 0 0 0 1 0 **0**
c	1 1 0 0 0 1 1 **1**
d	1 1 0 0 1 0 0 **0**
e	1 1 0 0 1 0 1 **1**
f	1 1 0 0 1 1 0 **1**
g	1 1 0 0 1 1 1 **0**
校验位	**0 0 1 1 1 1 1 0**

图4.4　水平垂直奇/偶校验

4.3.3 循环冗余校验码

循环冗余校验码(Cycle Redundancy Check,CRC)是一种被广泛采用的多项式编码,又称多项式码。循环冗余校验码具有良好的数学结构,易于实现,发送端编码器和接收端检测译码器的实现较为简单;同时,具有十分强的检错能力,特别适合于检测突发性的错误,在计算机网络中得到了广泛的应用。

CRC 码由两部分组成,前一部分是 $k+1$ 个比特的待发送信息,后一部分是 r 个比特的冗余码。由于前一部分是实际要传输的内容,因此是固定不变的,CRC 码的产生关键在于后一部分冗余码的计算。

计算中主要用到两个多项式:$f(x)$和 $G(x)$。其中,$f(x)$是一个 k 阶多项式,其系数是待发送的 $k+1$ 个比特序列;$G(x)$是一个 r 阶的生成多项式,由发收双方预先约定。例如,设要发送的信息序列是 1010001101(10 个比特,$k=9$),则以它们作为 $f(x)$的系数,得到对应的 9 阶多项式为

$$f(x)=1\times x^9+0\times x^8+1\times x^7+0\times x^6+0\times x^5+0\times x^4+1\times x^3+1\times x^2+0\times x+1$$
$$=x^9+x^7+x^3+x^2+1$$

再假设发收双方预先约定了一个 5 阶($r=5$)的生成多项式 $G(x)=x^5+x^4+x^2+1=1\times x^5+1\times x^4+0\times x^3+1\times x^2+0\times x+1$,则其系数序列为 110101。

CRC 码的产生方法如下。

(1) 生成 r 个比特的冗余码:用模 2 除法进行 $x^r f(x)/G(x)$运算,得余式 $R(x)$,其系数即是冗余码。

例如,$x^5 f(x)=x^{14}+x^{12}+x^8+x^7+x^5$,对应的二进制序列为 101000110100000,也就是 $f(x)$信息序列向左移动 $r=5$b,低位补 0。

$x^5 f(x)/G(x)=(101000110100000)/(110101)$,得余数为 01110,也就是冗余码,对应的余式 $R(x)=0\times x^4+x^3+x^2+x+0\times x^0$。

注意:① 若 $G(x)$为 r 阶,则 $R(x)$对应的比特序列长度为 r。

② 模 2 除法在做减法时不借位,相当于在进行异或运算。

(2) 得到带 CRC 校验的发送序列:用模 2 减法进行 $x^5 f(x)-R(x)$运算得到带 CRC 校验的发送序列,即 $x^5 f(x)-R(x)=101000110101110$。从形式上看,也就是简单地在原信息序列后面附加上冗余码。

在接收方,用同样的生成多项式 $G(x)$除所收到的序列。若余数为 0,则表示传输无差错,否则说明传输过程出现差错。例如,若收到的序列是 101000110101110,则用它除以同样的生成多项式 $G(x)=x^5+x^4+x^2+1$(即 110101)。因为所得余数为 0,所以收到的序列无差错。

CRC 校验方法是由多个数学公式、定理和推论得出的,尤其是 CRC 中的生成多项式对于 CRC 的检错能力会产生很大的影响,生成多项式 $G(x)$的结构及检错效果是在经过严格的数学分析和实验后才确定的,有其国际标准。常见的标准生成多项式如下。

CRC-12:$G(x)=x^{12}+x^{11}+x^3+x^2+1$

CRC-16:$G(x)=x^{16}+x^{15}+x^2+1$

CRC-32：$G(x)=x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+x^{8}+x^{7}$
$+x^{5}+x^{4}+x^{2}+x+1$

可以看出，只要选择足够的冗余位，就可以使得漏检率减少到任意小的限度。CRC 能够检验出下列差错。

（1）全部的奇数个错。

（2）全部的两位错。

（3）全部长度小于或等于 r 位的突发错。其中，r 是冗余码的长度。

由于 CRC 码的检错能力强，且容易实现，因此是目前应用非常广泛的检错码编码方法之一。CRC 码的生成和校验过程可以用软件或硬件方法来实现，如可以用移位寄存器和半加法器方便地实现。

例：若要传输的比特序列为 110011010，生成多项式 $G(x)=x^4+x^3+1$，求 CRC 码的校验序列码以及实际发送的比特序列是什么？

解析：发送数据对应的比特序列为 110011010；

生成多项式对应的比特序列为 11001（$k=4$）；

$F(x)\cdot x^k$ 对应的比特序列为 1100110100000；

将 $F(x)\cdot x^k$ 对应的比特序列用生成多项式对应的比特序列 11001 去除，计算过程及按二进制模 2 算法结果如图 4.5 所示；

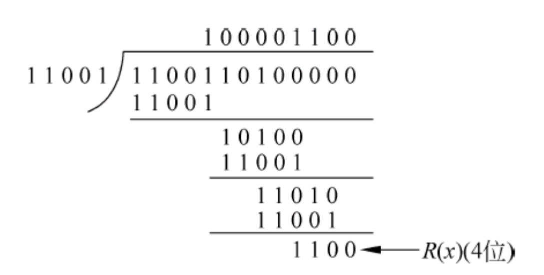


图 4.5　CRC 码的计算过程实例

求得余数比特序列为 1100（4b），就是 CRC 码的校验序列码；

实际发送的比特序列为 1100110101100。

例：若要传输的信息序列为 1000100101，生成多项式为 $G(x)=x^5+x^4+x^2+1$，求 CRC 码的校验序列码，若收到的数据比特序列为 100010010100011，验证其正确性。

解析：发送数据对应的比特序列为 1000100101；

生成多项式对应的比特序列为 110101（$k=5$）；

$F(x)\cdot x^k$ 对应的比特序列为 100010010100000；

将 $F(x)\cdot x^k$ 对应的比特序列用生成多项式对应的比特序列 110101 去除，计算过程及按二进制模 2 算法结果如图 4.6 所示；

求得余数比特序列为 00011（5b），就是 CRC 码的校验序列码；

将收到的比特序列用生成多项式对应的比特序列 110101 去除，若能够被整除，数据传输正确，否则传输错误，计算过程按二进制模 2 算法与前面相同，由读者自己完成。传输结果应是正确的。

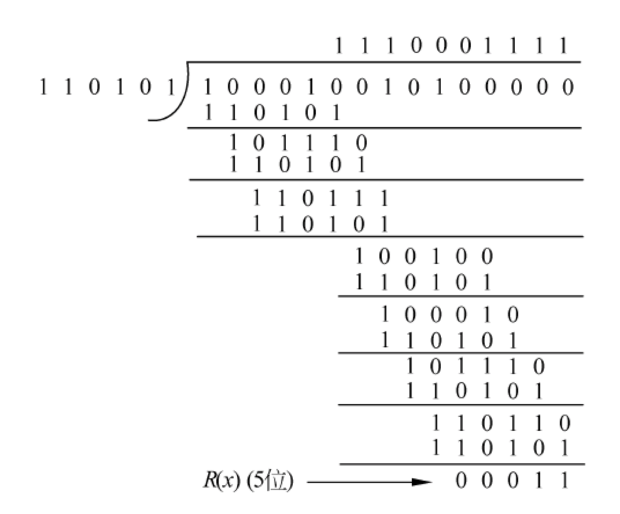


图 4.6　CRC 码的计算过程实例

4.3.4　反馈重发机制

由于检错码本身不提供自动的错误纠正能力，所以需要提供一种与之相配套的错误纠正机制，即反馈重发。通常当接收方检出错误的帧时，首先将该帧丢弃，然后给发送方反馈信息请求发送方重发相应的帧。反馈重发又被称为自动请求重传(Automatic Repeat Request，ARQ)。反馈重发有两种常见的实现方法，即停止—等待方式和连续 ARQ 方式。

1. 停止—等待方式

停止—等待方式(Stop-and-Wait)，也称为停—等协议。发送端在发出一帧之后必须停下来等待接收端的对发送帧的确认。若确认提示对方已经正确收到，则发送方继续发送下一个帧；否则，发送方就重发该帧。帧确认有肯定和否定之分，表示正确接收的被称为确认帧(Acknowledgement，ACK)，表示错误接收的被称为否认帧(Negative Acknowledgement，NAK)。

在理想情况下，帧在线路上不会损坏，也不会丢失。图 4.7 表明了发送方和接收方在正常情况下的一次帧传送。发送方发送一个帧，接收方在正确接收后，反馈一个确认帧。

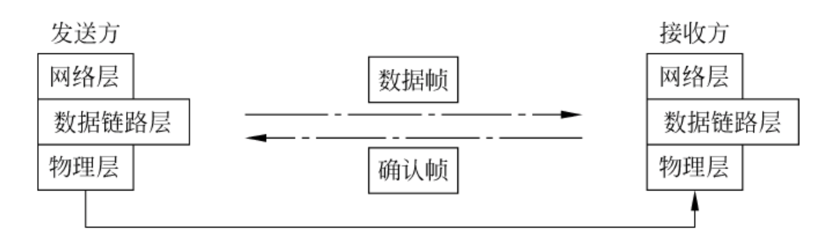


图 4.7　正常情况下的停止—等待

在这种理想的情况下，发送站点的动作如下。

(1) 将从网络层下来的分组装配成帧。

(2) 将帧送到数据链路层的发送缓冲区。

(3) 将发送缓冲区中的帧发送出去。

(4) 等待对方的确认。

(5) 收到接收站点的 ACK 帧,转到(1)。

接收站点的动作如下。

(1) 等待。

(2) 若收到由发送站点发过来的帧,将其放入数据链路层的接收缓冲区。

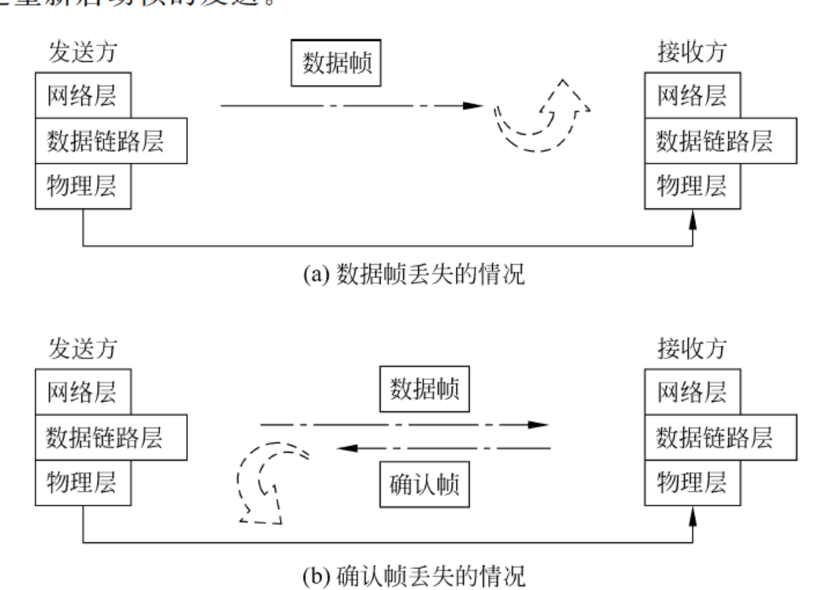
(3) 将接收缓冲区中的帧进行拆封处理,将其中的分组交给网络层。

(4) 发送一个确认帧 ACK 给发送站点,表示帧已正确接收。

(5) 转到(1)。

然而,任何信道都可能存在噪声。在有噪声情况下,帧可能被损坏,也可能完全丢失。有三种典型的帧丢失或损坏情况：一是接收方收到了被损坏的数据帧；二是发送端所发送的数据帧在传输过程被丢失；三是接收端发送给发送方的确认帧被丢失。在第一种情况下,接收方会发送一个否认帧,发送方在收到该否认帧后将会重传这个被损坏的帧。后两种情况则可能造成发送方无限制地等待下去。图 4.8 分别给出了后两种情况的示意。在图 4.8(a)中,由于接收方没有收到相应的帧,因此也就不会发送一个 ACK。在图 4.8(b)中,尽管接收方正确收到一个帧且为此发送了一个确认帧,但是该确认帧在传输过程中被丢失。所以,这两种情况都会造成发送方因收不到确认帧而无限等下去。

解决上述无限等待的有效方法是引入超时重发机制。在发送方设置一个计时器,当发送一个帧之后,就开始计时；如果在规定的时间内确认帧还未到达,就默认为帧在传输过程中被丢失,于是重新启动帧的发送。



(a) 数据帧丢失的情况

(b) 确认帧丢失的情况

图 4.8 两种帧丢失情况

但是,简单的超时重发会引发帧被重复接收的问题。例如,在图 4.8(b)中,数据帧已经被接收端正确接收,但接收端反馈的确认帧却在传输过程中丢失了,发送端因此启动超时重发机制,从而造成接收端收到重复的帧。解决帧重复接收问题的一个简单方法就是对待发送的帧进行编号。接收端一旦在某段时间内收到两个序列号相同的帧,就可以判断出它们是重复的帧,然后丢弃重复的帧。停止—等待方式实现简单,但是这种发送一帧等待一个确认的方式使得通信效率很低。为此,人们提出了连续 ARQ 方式。

2. 连续 ARQ 方式

连续重发请求方案是指发送方可以连续发送一系列信息帧，即不用等前一帧被确认便可发送下一帧。这就需要在发送方设置一个较大的缓冲存储空间(称为重发表)，用以存放若干待确认的信息帧。当发送方收到对某信息帧的确认帧后便可从重发表中将该信息帧删除。所以，连续 ARQ 方式的链路传输速率大大提高，但相应地需要更大的缓冲存储空间。连续 ARQ 方式的实现过程如下。

(1) 发送方连续发送信息帧而不必等待确认帧的返回。

(2) 发送方在重发表中保存所发送的每个帧的备份。

(3) 重发表按先进先出(FIFO)队列规则操作。

(4) 接收方对每一个正确收到的信息帧返回一个确认帧。

(5) 每一个确认帧包含一个唯一的序号，随相应的确认帧返回。

(6) 接收方保存一个接收次序表，它包含最后正确收到的信息帧的序号。

(7) 当发送方收到相应信息帧的确认后，从重发表中删除该信息帧的备份。

(8) 当发送方检测出失序的确认帧(即第 N 号信息帧和第 $N+2$ 号信息帧的确认帧已返回，而第 $N+1$ 号的确认帧未返回)后，便重发未被确认的信息帧。

上面连续 ARQ 过程是假定在不发生传输差错的情况下描述的，如果差错出现，如何进一步处理还可以有两种策略，即 Go-Back-N 策略和选择重发策略。

Go-Back-N 策略的基本原理是，当接收方检测出失序的信息帧后，要求发送方重发最后一个正确接收的信息帧之后的所有未被确认的帧；或者当发送方发送了 N 个帧后，若发现该 N 帧的前一个帧在计时器超时后仍未返回其确认信息，则该帧被判为出错或丢失，此时发送方就不得不重新发送出错帧及其后的 N 帧。这就是 Go-Back-N(退回 N)策略名称的由来。因为，对接收方来说，由于这一帧出错，就不能以正常的序号向它的高层递交数据，对其后发送来的 N 帧也可能都不能接收而丢弃。Go-Back-N 策略操作过程如图 4.9 所示。图中假定发送完 8 号帧后，发现 2 号帧的确认返回在计时器超时后还未收到，则发送方只能退回从 2 号帧开始重发。

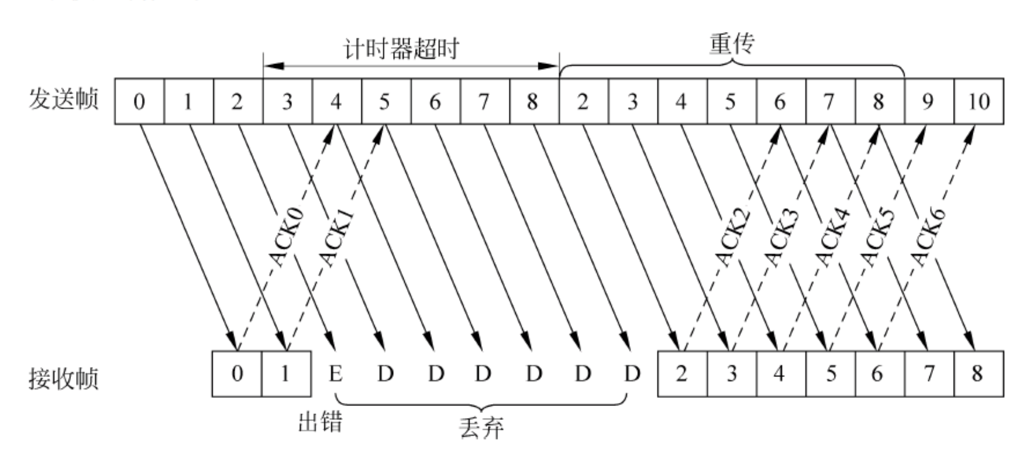


图 4.9 Go-Back-N 举例

Go-Back-N 可能将已正确传送到目的方的帧再重传一遍，这显然是一种浪费。另一种效率更高的策略是当接收方发现某帧出错后，其后继续送来的正确的帧虽然不能立即递交给接收方的高层，但接收方仍可收下来，存放在一个缓冲区中，同时要求发送方重新传送出

错的那一帧。一旦收到重新传来的帧后，就可以将原已存于缓冲区中的其余帧一并按正确的顺序递交高层。这种方法称为选择重发（Selective Repeat），其工作过程如图 4.10 所示。图中 2 号帧的否认返回信息 NAK2 要求发送方选择重发 2 号帧。显然，选择重发减少了浪费，但要求接收方有足够大的缓冲区空间。

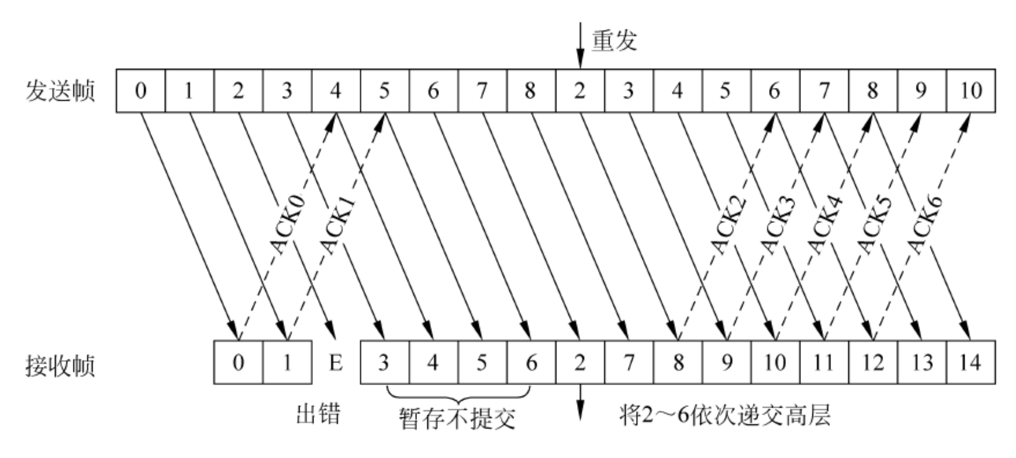


图 4.10 选择重发举例法

4.4 流量控制

4.4.1 流量控制作用

由于系统性能的不同，如硬件能力（包括 CPU、存储器等）和软件功能的差异，会导致发送方与接收方处理数据的速度有所不同。若一个发送能力较强的发送方给一个接收能力相对较弱的接收方发送数据，则接收方会因无能力处理所有收到的帧而不得不丢弃一些帧。如果发送方持续高速地发送，则接收方最终会被"淹没"。也就是说，在数据链路层只有差错控制机制还是不够的，其不能解决因发送方和接收方速率不匹配所造成的帧丢失。为此，在数据链路层引入了流量控制机制。流量控制并不是数据链路层特有的功能，许多高层协议中也提供流量控制功能，只不过流量控制的对象不同而已。比如，对于数据链路层来说，控制的是相邻两节点之间数据链路上的流量；而对于运输层来说，控制的则是从源到最终目的之间端对端的流量。

流量控制的作用就是使发送方所发出的数据流量不要超过接收方所能接收的速率。流量控制的关键是需要有一种信息反馈机制，使发送方能了解接收方是否具备足够的接收及处理能力。存在各种不同的流量控制机制。如上面所提到的简单停—等协议就可以实现流量控制功能，但其实现效率太低。下面所介绍的滑动窗口协议可以将流量控制机制与帧确认机制巧妙地结合在一起。

4.4.2 滑动窗口协议

为了提高信道的有效利用率，如前所述采用了不等待确认帧返回就连续发送若干帧的

方案。由于允许连续发送多个未被确认的帧，帧号就需采用多位二进制才能加以区分。因为凡被发出去但尚未被确认的帧都可能出错或丢失而要求重发，因而这些帧都要保留下来。这就要求发送方有较大的发送缓冲区保留可能要求重发的未被确认的帧。

但是缓冲区容量总是有限的，如果接收方不能以发送方的发送速率处理接收到的帧，则还是可能用完缓冲容量而暂时过载。为此，可引入类似于空闲 ARQ 控制方案的调整措施，其本质是在收到一确定帧之前，对发送方可发送的帧的数目加以限制。这是由发送方调整保留在重发表中的待确认帧的数目来实现的。如果接收方来不及对新到的帧进行处理，则便停发确认信息，此时发送方的重发表就会增长，当达到重发表限度时，发送方就不再发送新帧，直至再次收到确认信息为止。

为了实现此方案，发送方存放待确认帧的重发表中，应设置待确认帧数目的最大限度，这一限度被称为链路的发送窗口。显然，如果窗口设置为 1，即发送方缓冲能力仅为一个帧，则传输控制方案就回到了停—等协议，此时传输速率很低。故窗口限度应选为使接收方尽量能处理或接收收到的所有帧。当然选择时还必须考虑诸如帧的最大长度、可使用的缓冲区空间以及传输速率等因素。

重发表是一个连续序号的列表，对应发送方已发送但尚未确认的那些帧。这些帧的序号有一个最大值，这个最大值即发送窗口的限度。所谓发送窗口就是指发送方已发送但尚未确认的帧序号队列的界，其上下界分别称为发送窗口的上下沿，上下沿的间距称为窗口尺寸。接收方类似地也有接收窗口，它指示允许接收的帧的序号。

发送方每次发送一帧后，待确认帧的数目便增 1；每收到一个确认信息后，待确认帧的数目便减 1。当重发表长度计数值，即待确认帧的数目等于发送窗口尺寸时，便停止发送新的帧。

一般帧号只取有限位二进制数，到一定时间后就又反复循环。若帧号配 3b 二进制数，则帧号在 0～7 间循环。如果发送窗口尺寸取值为 2，则发送如图 4.11 所示。图中发送方阴影部分表示打开的发送窗口，接收方阴影部分则表示打开的接收窗口。当传送过程进行时，打开的窗口位置一直在滑动，所以也称为滑动窗口（Sliding Window），或简称为滑窗。

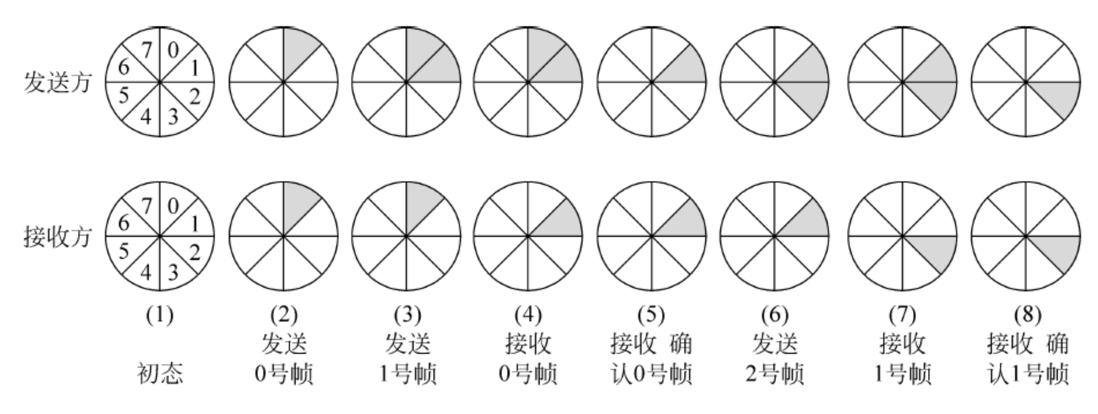


图 4.11 滑动窗口状态变化过程

图 4.11 中的滑动窗口状态变化过程可叙述如下（假设发送窗口尺寸为 2，接收窗口尺寸为 1）。

(1) 初始态，发送方没有帧发出，发送窗口前后沿相重合。接收方 0 号窗口打开，表示

等待接收0号帧。

(2) 发送方已发送0号帧,此时发送方打开0号窗口,表示已发出0号帧但尚未收到确认返回信息。此时接收窗口状态同前,仍等待接收0号帧。

(3) 发送方在未收到0号帧的确认返回信息前,继续发送1号帧。此时,1号窗口打开,表示1号帧也属等待确认之列。至此,发送方打开的窗口数已达规定限度,在未收到新的确认返回帧之前,发送方将暂停发送新的数据帧。接收窗口此时状态仍未变。

(4) 接收方已收到0号帧,0号窗口关闭,1号窗口打开,表示准备接收1号帧。此时发送窗口状态不变。

(5) 发送方收到接收方发来的0号帧确认返回信息,关闭0号窗口,表示从重发表中删除0号帧。此时接收窗口状态仍不变。

(6) 发送方继续发送2号帧,2号窗口打开,表示2号帧也纳入待确认之列。至此,发送方打开的窗口又已达规定限度,在未收到新的确认返回帧之前,发送方将暂停发送新的数据帧,此时接收窗口状态仍不变。

(7) 接收方已收到1号帧,1号窗口关闭,2号窗口打开,表示准备接收2号帧。此时发送窗口状态不变。

(8) 发送方收到接收方发来的1号帧收毕的确认信息,关闭1号窗口,表示从重发表中删除1号帧。此时接收窗口状态仍不变。

一般来说,凡是在一定范围内到达的帧,即使它们不按顺序,接收方也要接收下来。若把这个范围看成是接收窗口,则它的大小也应该是大于1的。而Go-Back-*N*正是接收窗口等于1的一个特例,选择重发也可以看作是一种滑动窗口协议,只不过其发送窗口和接收窗口都大于1。若从滑动窗口的观点来统一看待停—等协议、Go-Back-*N*策略及选择重发策略三种协议,它们的差别仅在于各自窗口尺寸的大小不同而已:

(1) 停—等协议:发送窗口=1,接收窗口=1;

(2) Go-Back-*N*策略:发送窗口>1,接收窗口=1;

(3) 选择重发策略:发送窗口>1,接收窗口>1。

4.5 数据链路层协议示例

4.5.1 HDLC——高级数据链路控制

高级数据链路控制(规程)(High Level Data Link Control,HDLC)是一个在同步网上传输数据、面向位的数据链路层协议,它是由国际标准化组织(ISO)制定的。HDLC是IBM的同步数据链路控制规程(SDLC)的一个超集。

HDLC是面向比特的协议,支持全双工通信,采用位填充的成帧技术,以滑动窗口协议进行流量控制。

1. HDLC的帧格式

HDLC的功能集中体现在HDLC帧格式中,HDLC的帧格式如图4.12所示。

(1) 帧头和帧尾的位模式串01111110为帧的开始和结束标记(Flag)。可以看出,

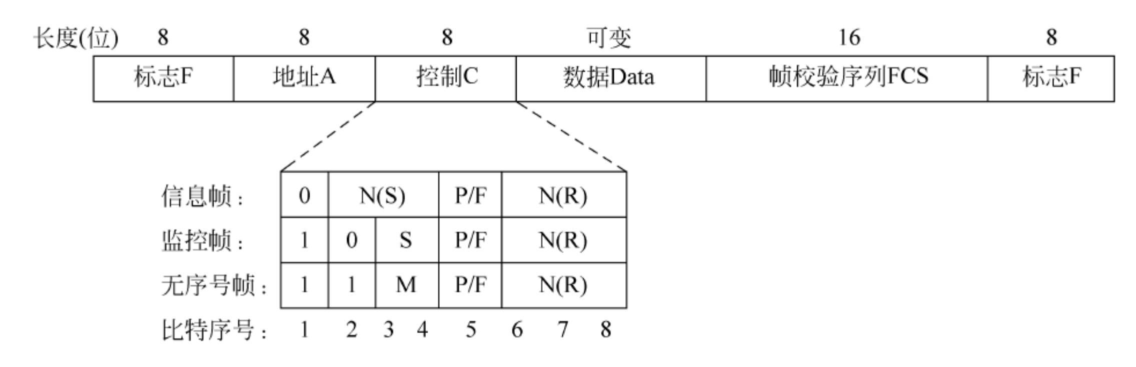


图 4.12 HDLC 的帧格式及控制字段的结构

HDLC 协议在帧定界上采用的是用比特填充的首尾定界符法。

(2) A 是地址字段(Address),由 8b 组成。对于命令帧,存放接收站的地址;对于响应帧,存放发送响应帧的站点地址。

(3) C 是控制字段(Control),由 8b 组成,该字段是 HDLC 协议的关键部分。它标志了 HDLC 的 3 种类型帧:信息(Information)帧、监控(Supervisory)帧和无序号(Unnumbered)帧。如图 4.12 中关于控制字段结构所示,若帧的第 1 比特为 0,则代表这是一个用于发送数据的信息帧,相应地,其第 2～4 比特代表当前发送的信息帧的序号,而第 6～8 比特则代表接收序号即期望收到的帧的发送序号。若帧的第 1 比特和第 2 比特为 10,则代表这是一个用于协调双方通信状态的监控帧,相应地,其第 3 比特和第 4 比特用以代表 4 种不同类型的监控帧。00 表示接收准备就绪;01 表示传输出错,并要求采用拉回方式重发;10 表示接收准备尚未就绪,要求发送方暂停发送;11 则表示传输出错并要求采用选择重发。

(4) 监控帧中不包含 Data(数据)部分,若帧的第 1 比特和第 2 比特为 11,则代表用于数据链路控制的无序号帧,其第 3 比特、第 4 比特、第 6 比特、第 7 比特和第 8 比特用 M(Modifier)表示,M 的取值不同表示不同功能的无序号帧。无序号帧可用于建立连接和拆除连接。在所有 3 种情况下,第 5 比特是轮询/终止(Poll/Final)比特,简称 P/F,用于询问对方是否有数据要发送或告诉对方数据传输结束。

(5) Data 是数据字段,可以包含任意信息且可以是任意长的,但实际上受多种条件的制约,如帧校验效率就会随着数据长度的增加而下降。

(6) FCS 是校验序列字段,采用 16b 的 CRC 校验,其生成多项式为 CRC－16:

$G(x)=x^{16}+x^{12}+x^5+1$,校验的内容包括 A 字段、C 字段和 Data 字段。

2. HDLC 的帧类型

HDLC 提供了 3 种类型的帧,分别是信息帧 I、监控帧 S 和无序号帧 U。一个 HDLC 帧属于哪种类型取决于控制字段中的第 1 位或前 2 位。对于这 3 种类型的帧,控制字段中的第 5 比特都是轮询/终止比特,用于询问对方是否有数据要发送或告诉对方数据传输结束。但是该比特的功能较多,在不同帧中用法也不一样。

1) 信息帧

信息帧简称 I 帧,其控制字段的第 1 比特为 0。I 帧用于发送数据。在 I 帧中,第 2～4 比特 N(S)代表当前发送的 I 帧的序号。第 6～8 比特 N(R)则代表接收序号,它给出下一个期望接收的帧序号。也就是说,接收方不必专门为正确收到的 I 帧发送确认,它可以在自己当前所要发送的 I 帧中通过 N(R)捎带上对已经正确接收帧的确认。例如,在连续收到对方

N(S)=0～3 的 4 个 I 帧后，可将 N(R)置为 4，表示 3 号帧及其以前的各帧均已经正确收到，而下一个期望接收的是发送序号 N(S)=4 的 I 帧。这种捎带(Piggyback)技术可以提高信道的利用率。

应该注意，由于通信是全双工的，所以参与通信的每一方各有一个 N(S)和一个 N(R)。另外，对于 I 帧中的 P/F 位，若 P/F=0，表示该位没有意义；若 P/F=1，则对不同的情况有不同的含义。对于正常响应模式 NRM，若主站置 P/F=1，则表示它向从站发出探询，只有从站有数据要发送，才可以向主站发送信息。当从站发送最后一帧时，它要置 P/F=1，表示从站数据发送已经结束。

2）监控帧

监控帧简称 S 帧，其控制字段的第 1 比特为 1、第 2 比特为 0。S 帧用于协调双方通信状态。根据 S 帧中控制字段的第 3 比特至第 4 比特取值，S 帧又进一步分成 4 种类型，这 4 种 S 帧的名称及作用如表 4.1 所示。

表 4.1　4 种类型的 S 帧

第 3 比特与 第 4 比特的取值	帧 名 称	功　　能
00	接收准备就绪 (Receive Ready,RR)	用于连续 ARQ 中。表示确认序号为 N(R)-1 及其以前的各帧，等待接收序号为 N(R)的帧。具有流量控制功能
01	拒绝接收 (REJect,REJ)	用于连续 ARQ 中。表示传输出错，N(R)以后的各帧被否认，并要求采用拉回方式重发。但确认 N(R)-1 及其以前的各帧
10	接收尚未就绪 (Receive Not Ready,RNR)	用于连续 ARQ 中。表示暂停接收下一帧，但确认 N(R)-1 及其以前的各帧。具有流量控制功能
11	选择拒绝 (Selective REJect,SREJ)	用于选择式 ARQ 中。否认序号为 N(R)的帧，并要求采用选择重发，但确认 N(R)-1 及其以前的各帧

所有 S 帧都不包含要传送的数据信息，所以它不需要有发送序号 N(S)。但是，S 帧中的接收序号 N(R)却仍然非常重要。在 RR 帧和 RNR 帧中，N(R)都相当于是对前面已经正确接收的各帧的确认，N(R)表示下一个期望接收的帧；在 REJ 监控帧中，N(R)表示否认的信息帧号，同时也是对 N(R)-1 及其以前的各帧均已正确接收的一种确认。

对于 S 帧，当 P/F 位的取值为 0 时也是没有意义的。P/F 位取 1 才有意义。在非平衡配置的正常响应方式中，从站不能主动向主站发送信息，只有收到主站发来的 P/F=1 的 S 帧(或 I 帧)之后才能发送响应帧。若从站有数据要发送，则在最后一个数据帧中将 P/F 位置 1。若无数据发送，则在回答的 S 帧中将 P/F 位置 1。

在非平衡配置的异步响应模式或在平衡配置的异步平衡模式中，因为任何一个站都可以主动向对方发送数据，所以在主动发送的 S 帧和 I 帧中都可将 P/F 位置 1，接收方收到 P/F 位为 1 的帧后，就尽快地回答本站的状态并将 P/F 比特置为 1。

3）无序号帧

无序号帧简称 U 帧，其控制字段的第 1 比特和第 2 比特为 11。无序号帧本身不带编号，即无 N(S)和 N(R)字段，其第 3、第 4、第 6、第 7 和第 8 比特用 M 表示，M 的取值不同表示不同功能的无序号帧。虽然共有 32 个不同的序号，但目前只定义了 15 种无序号帧。无序号帧主要用于数据链路控制，如用于链路连接的建立和拆除。表 4.2 给出了 6 种 U 帧控制字段(C 字段)的第 3～8 比特的格式与链路控制功能。在这 6 种 U 帧中，前 4 种是命令，后 2 种是从站的响应。

表 4.2　6 种用于数据链路控制的 U 帧格式与功能

U 帧名	命令/响应	M		P/F	M			功　　能
置异步响应	SARM，命令	1	1		0	0	0	建立主从的点对点结构
置正常响应	SNRM，命令	0	0		0	0	1	建立主从的多点结构
置异步平衡响应	SABM，命令	1	1		1	0	0	建立复合站的平衡结构
拆链	DISC，命令	0	0		0	1	0	结束已建立的数据链路
无序号确认	UA，响应	0	0		1	1	0	从站响应主站的命令
命令拒绝	CMAD，响应	1	0		0	0	1	从站报告帧传输异常

3. HDLC 用于实现面向连接的可靠传输

图 4.13 给出了将 HDLC 用于实现有确认的面向连接数据传输服务的例子。图 4.13 为正常传输，其中将无序号帧用于数据链路连接的建立、维护与拆除，而信息帧用于发送数据并实现捎带的帧确认。图 4.14 则表示出现差错后的处理过程，但省略了关于连接建立的过程。由于 B 方没有数据帧要发送给 A 方，所以不能利用信息帧的捎带来反馈帧出错信息，只有专门发送一个监控帧用于告诉 A 方数据帧传输出错并同时给出建议的差错控制方式，显然在该例子中差错控制采用了选择重发方式。

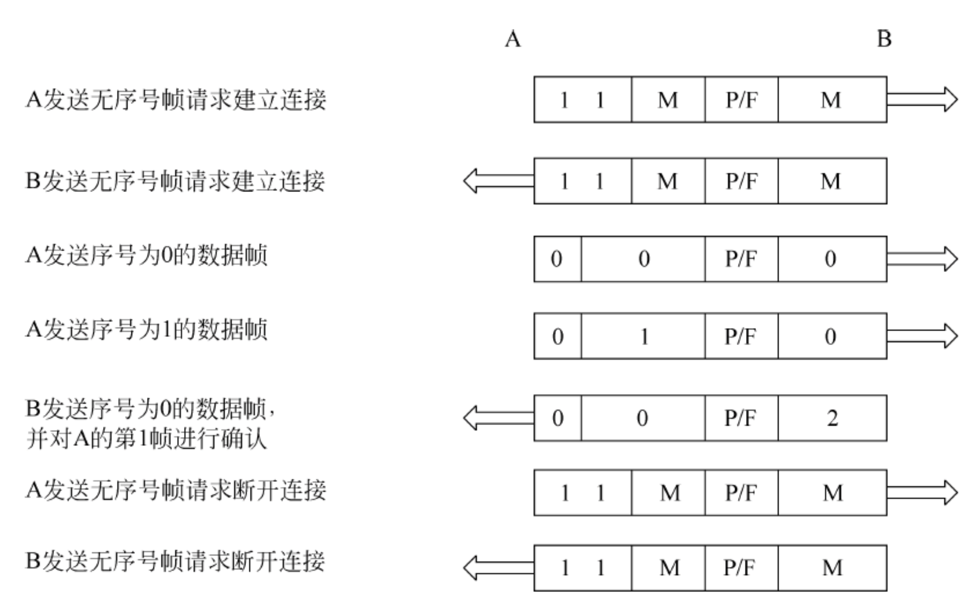


图 4.13　有确认的面向连接 HDLC 的连接建立、数据传输和连接拆除

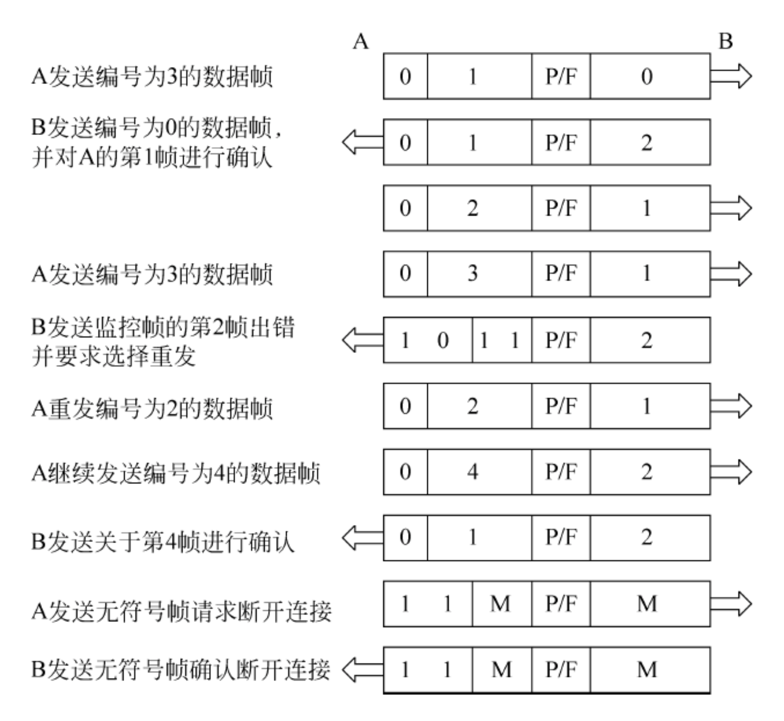


图 4.14 有确认的面向连接 HDLC 差错控制的实现连接建立、数据传输和连接拆除

4.5.2 PPP

PPP 是点对点协议(Point-to-Point Protocol)的简称，它是一个工作于数据链路层的广域网协议。PPP 由 IETF(Internet Engineering Task Force)开发，目前已被广泛使用并成为国际标准。无论是同步电路还是异步电路，PPP 都能够建立路由器之间或者主机到网络之间的连接，如图 4.15 所示。例如利用 Modem 进行拨号上网(163、169、165 等)就是使用 PPP 实现主机到网络连接的典型例子。

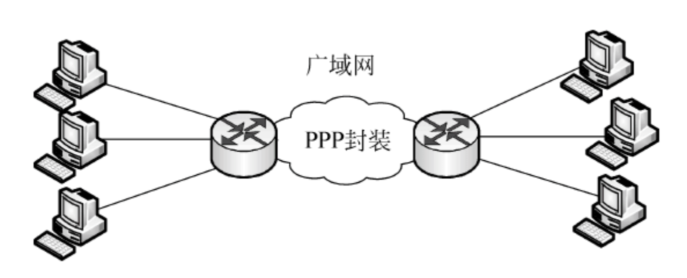


图 4.15 PPP 提供多种连接

1. PPP 的特性

PPP 是目前使用最广泛的广域网协议，这是因为它具有以下特性。

(1) 能够控制数据链路的建立；

(2) 能够对 IP 地址进行分配和使用；

(3) 允许同时采用多种网络层协议；

(4) 能够配置和测试数据链路；

(5) 能够进行错误检测；

(6) 有协商选项,能够对网络层的地址和数据压缩等进行协商。

PPP 是现在主流的一种国际标准 WAN 封装协议,可支持如下连接类型。

(1) 同步串行连接;

(2) 异步串行连接;

(3) ISDN 连接;

(4) HSSI 连接。

2. PPP 的组成

PPP 作为第 2 层的协议,在物理上可使用各种不同的传输介质,包括双绞线、光纤及无线传输介质,在数据链路层提供了一套解决链路建立、维护、拆除和上层协议协商、认证等问题的方案;在帧的封装格式上,PPP 采用的是一种 HDLC 的变化形式;其对网络层协议的支持则包括了多种不同的主流协议,如 IP 和 IPX 等。图 4.16 给出了 PPP 的体系结构,其中,链路控制协议(Link Control Protocol,LCP)用于数据链路连接的建立、配置与测试,NCP(Network Control Protocol)则是一组用来建立和配置不同数据链路的网络层协议。

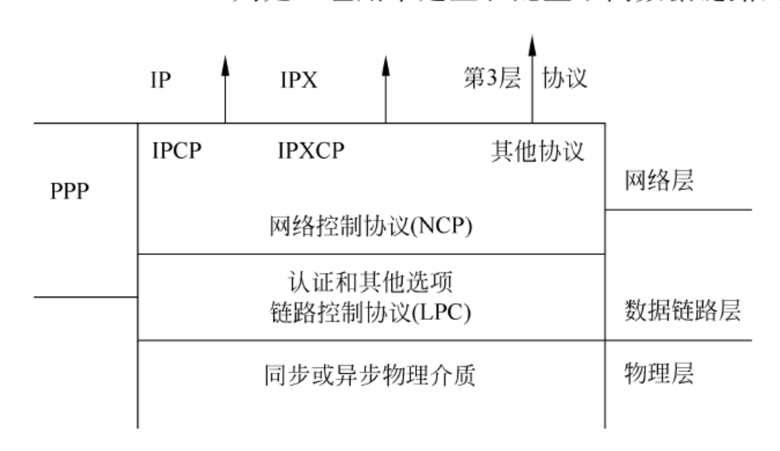


图 4.16　PPP 结构

3. PPP 会话建立的过程

PPP 提供了建立、配置、维护和终止点到点连接的方法。PPP 经过以下 4 个阶段在一个点到点的链路上建立通信连接。

(1) 链路的建立和配置协调:通信的发起方发送 LCP 帧来配置和检测数据链路。LCP 帧有链路建立帧、链路终止帧和链路维护帧 3 种。在链路建立阶段主要是通过发送 LCP 的帧来对链路进行相关的配置,包括数据的最大传输单元、是否采用 PPP 的压缩、PPP 的认证方式等。

(2) 链路质量检测:在链路建立、协调之后。这一阶段是可选的。主要用于对链路质量进行测试,以确定其能否为上层所选定的网络协议提供足够的支持,另外,若连接的双方已经要求采用安全认证,则在该阶段还要按所选定的认证方式进行相应的身份认证。

(3) 网络层协议配置协调:通信的发起方发送 NCP 帧以选择并配置网络层协议。配置完成后,通信双方可以发送各自的网络层协议数据报。通过发送 NCP 帧来选择网络层协议并进行相应的配置,不同的网络层协议要分别进行配置。此时,一条完整的 PPP 链路就建立起来了,可在所建立的 PPP 链路上进行数据传输。

(4) 关闭链路:通信链路将一直保持到 LCP 或 NCP 帧关闭链路或者是发生一些外部

事件(如空闲时间超长或用户干预)。

需要说明的是,尽管 PPP 的验证是一个可选项,但一旦采用身份验证,则必须在网络层协议阶段之前进行。有两种类型的 PPP 验证,即 PAP(Password Authentication Protocol)与 CHAP(Challenge Handshake Authentication Protocol)方式。PAP 采用的是两次握手方式,远程节点提供用户名与密码,由本地节点提供身份验证的确认或拒绝。

用户名与密码对由远程网络节点不断地在链路上发送,直到验证被确认或被终结。密码在传输过程中采用的是明文方式,而且发送登录请求的时间和频率完全由远程节点控制,所以这种验证方式虽然实现简单但易受到攻击。CHAP 所使用的是三次握手的验证方式,本地节点提供一个用于身份验证的挑战值,由远程节点根据所收到的挑战值计算出一个回应值发送回本地节点,若该值与本地节点的计算结果一致,则远程节点被验证通过。显然一个没有获得挑战值的远程节点是不可能尝试登录并建立连接的,也就是说,CHAP 由本地来控制登录的时间与频率,并且由于每次所发送的挑战值都是一个不可预测的随机变量,所以 CHAP 较之 PAP 更加安全有效,因此在通常情况下,更多采用的是 CHAP 验证方式。

课后习题

1. 术语解释

差错　纠错码　检错码　流量控制　差错控制　帧　面向无连接的服务　面向连接的服务

2. 数据链路(即逻辑链路)与链路(即物理链路)有何区别?“电路接通了”与“数据链路接通了”的区别何在?

3. 数据链路层的主要任务和功能是什么?

4. 数据链路层所提供的基本服务有哪些?

5. 在数据帧中,当所传的数据中出现了控制字符时,就必须采取适当的措施,使接收方不至于将数据误认为是控制信息,这样才能保证数据链路层的传输是(　　)的。

A. 透明　　B. 面向连接　　C. 冗余　　D. 无连接

6. 数据链路层服务功能主要可以分为三类:面向连接确认服务、无连接确认服务和(　　)。

A. 差错控制服务　　B. 面向连接不确认服务
C. 认证服务　　D. 无连接不确认服务

7. 在(　　)差错控制方式中,只会重新传输出错的数据帧。

A. 连续工作　　B. 停止—等待　　C. 选择重发　　D. 拉回

8. 一个信道速率为 4Kbps,采用停止—等待协议,传播时延为 20ms,确认帧长度和处理时间均可忽略,当帧长为多少才能使信道利用率达到至少 50%?

9. 若要传送的信息序列为 1000100101,生成多项式为 $G(x)=x^5+x^4+x^2+1$,求 CRC 码的校验序列码,若收到的数据比特序列为 100010010100011,验证其正确性。

10. 何谓差错?引起差错的原因是什么?

11. 什么是成帧?数据链路层常用的成帧方法有哪些?

12. 试简述 HDLC 帧各字段的意义。HDLC 用什么方法保证数据的透明传输?

13. 在停止—等待协议中，确认帧是否需要序号？请说明理由。

14. 发送数据比特序列为 110011(6b)，生成多项式比特序列为 11001，求 CRC 校验序列。

15. HDLC 帧可分为哪几个大类？试简述各类帧的作用。

16. 简述滑动窗口的原理。

17. 要发送的数据为 1101011011，采用 CRC 的生成多项式是 $P(x)=x^4+x+1$，试求应添加在数据后面的余数。

A. 若数据在传输过程中最后一个 1 变成 0，问接收端能否发现？

B. 若数据在传输过程中最后两个 1 变成 0，问接收端能否发现？

C. 若采用 CRC 检验后，数据链路层的传输是否就变成了可靠的传输？

18. 要发送的数据为 101110，采用 CRC 的生成多项式是 $P(x)=x^3+1$，试求应添加在数据后面的余数。

第5章 局域网

学习目的

本章在介绍局域网的两个子层关键技术与标准的基础上，对共享介质局域网、高速与高速局域网的工作原理，以及城域网、无线局域网的工作原理进行了系统的讨论。

通过本章的学习，使读者能够了解局域网与城域网的主要技术特点，掌握 Ethernet 局域网，以及高速局域网、交换局域网、无线局域网以及虚拟局域网的基本工作原理，掌握局域网互联的基本概念，初步具备局域网组网的基础知识与能力。

学习要求

了解：局域网与城域网的主要技术特点。

了解：局域网拓扑结构的类型与特点。

了解：IEEE 802 参考模型与协议的基本概念。

掌握：Ethernet 局域网的基本工作原理。

了解：令牌环网与 FDDI 的基本工作原理。

掌握：高速局域网、交换局域网与虚拟局域网的基本工作原理。

了解：无线局域网的基本工作原理。

掌握：网桥的基本工作原理。

5.1 局域网概述

局域网是计算机网络的重要组成部分，是当今计算机网络技术应用与发展非常活跃的一个领域。公司、企业、政府部门及住宅小区内的计算机都通过 LAN 连接起来，以达到资源共享、信息传递和数据通信的目的。而信息化进程的加快，更是刺激了通过 LAN 进行网络互联需求的剧增。因此，理解和掌握局域网技术也就显得很重要。

局域网的发展始于 20 世纪 70 年代，至今仍是网络发展中的一个活跃领域。到了 90 年代，LAN 更是在速率、带宽等指标方面有了更大进展，并且在 LAN 的访问、服务、管理、安全和保密等方面都有了进一步的改善。例如，Ethernet 技术从传输速率为 10Mbps 的 Ethernet 发展到 100Mbps 的高速以太网，并继续提高至千兆位(1000Mbps)以太网、万兆位以太网。

5.1.1 局域网的主要特点与功能

局域网技术是当前计算机网络研究与应用的一个热点问题，也是目前技术发展非常迅

速的领域之一。局域网最主要的特点是：网络为一个单位所拥有，且地理范围和站点数目均有限。局域网具有如下特点。

(1) 网络所覆盖的地理范围比较小。通常不超过几十千米，甚至只在一个园区、一幢建筑或一个房间内。

(2) 数据的传输速率比较高，从最初的1Mbps到后来的10Mbps、100Mbps，近年来已达到1000Mbps、10 000Mbps。

(3) 局域网有较低的时延和较低的误码率。由于局域网采用专线连接，其信息传输就可以避免广域网传输中信号经过多次交换而产生的时延和干扰，这样信息在传输时就具备较低的时延和较低的误码率。

(4) 局域网的经营权和管理权属于某个单位所有，与广域网通常由服务提供商提供形成鲜明对照。

(5) 局域网一般采用广播技术而非交换技术，这是因为局域网中的通信是在共享传输媒体上进行的，所以在局域网中，各个站点能够进行广播(一站向其他所有站发送)或组播(一站向多个站发送)。

(6) 便于安装、维护和扩充，建网成本低、周期短。

正是由于局域网以上的特点，在LAN的设计过程中，其实现的关键技术为拓扑、传输媒体和媒体的访问控制协议。

应该指出，尽管局域网地理覆盖范围小，但这并不意味着它们必定是小型的或简单的网络。随着网络互联技术的发展与网络互联设备性能的提高，局域网可以扩展得相当大或者非常复杂，具有成千上万用户的局域网是很常见的事。局域网具有如下的一些主要优点。

(1) 能方便地共享昂贵的外部设备、主机以及软件、数据，从一个站点可访问全网。

(2) 便于系统的扩展和逐渐地演变，各设备的位置可灵活调整和改变。

(3) 提高了系统的可靠性、可用性。

局域网的功能概括起来就是资源共享、数据交换，具体可以分为以下几种功能。

1. 文件共享

人们在实际应用中可能都有这样的体会，如果没有网络，想给别人传输一些软件或文件，会非常不方便。现在的数据动辄就是几兆、几十兆、几百兆甚至更大，用闪存或移动硬盘虽然能解决传输问题，但在不同计算机之间插拔闪存或移动硬盘比较麻烦，刻成光盘又需要光盘刻录机，而使用局域网，文件的传输就轻松多了，100Mbps局域网传输文件的速率可以达到每秒1～3MB，相当于普通光驱的速率。只要将要传输的文件设为共享，传输是非常容易的。

2. 互联网接入共享

建立局域网之后，局域网内的计算机可以以共享连接的方式，统统接入互联网，可以浏览网页，可以收发邮件，可以在局域网内安全、快速、廉价地享受网上冲浪的乐趣。

3. 内部网站和BBS

有了局域网之后，可以架设在局域网内部使用的网站，供内部人员浏览；可以建立内部的BBS系统，大家可以在此交流、发布信息等。

4. 电子邮件

在局域网内，可以架设自己的电子邮件服务器，可以在不经过互联网的情况下实现内部

电子邮件的发送和接收,非常快捷;也可以收发来自互联网上的邮件。

5. 网络打印

给每台计算机配备打印机在当前的国情下是不太现实的。如果架设了局域网,就可以通过网络共享一台或若干台打印机,这样就会经济地实现打印资源的共享。

6. 基于网络的其他应用

以上仅是局域网的通用功能,是对局域网最基本的应用。针对每个局域网用户的实际情况,还可以通过局域网和相关硬件及软件系统,实现各种信息化。如可以实现"一卡通",只要安装必要的软件和硬件,就可以在局域网覆盖的地方持一个IC卡实现购物、借阅、餐饮等功能。在局域网的基础上,还可以实现各种办公自动化,如考勤、请假、报销、公文流转和报批、通知等。各行各业的信息化均有不同,这里就不一一详述局域网的其他特定用途了。

值得注意的是,局域网只是为信息化提供基础。企事业单位的信息化更多的是在网络应用软件方面,而不是局域网的基本硬件和软件。

5.1.2 局域网的组成

要构成LAN,必须有其基本部件。LAN既然是一种计算机网络,自然少不了计算机,特别是个人计算机(PC)。几乎没有一种网络只由大型机或小型机构成。因此,对于LAN而言,个人计算机是一种必不可少的构件。计算机互联在一起,当然也不可能没有传输媒体,这种媒体可以是同轴电缆、双绞线、光缆或辐射性媒体。第3个构件是任何一台独立计算机通常都不配备的网卡,也称为网络适配器,但在构成LAN时,则是不可少的部件。第4个构件是将计算机与传输媒体相连的各种连接设备,如DB-15插头座、RJ-45插头座等,具备了上述4种网络构件,便可搭成一个基本的LAN硬件平台。

有了LAN硬件环境,还需要控制和管理LAN正常运行的软件,即网络操作系统(Network Operate System,NOS),是在每台PC原有操作系统上增加网络所需的功能。例如,当需要在LAN上使用字处理程序时,用户的感觉犹如没有组成LAN一样,这正是LAN操作发挥了对字处理程序访问的管理。在LAN情况下,字处理程序的一个复制通常保存在文件服务器中,并由LAN上的任何一个用户共享。由上面介绍的情况可知,组成LAN需要下述5种基本结构:

(1) 计算机(特别是PC);

(2) 传输介质;

(3) 网络适配器;

(4) 网络连接设备;

(5) 网络操作系统。

5.1.3 局域网的拓扑结构与传输介质

1. 局域网的拓扑结构

局域网的拓扑结构主要有4种:总线型网、环形网、星形网、树形网,如图5.1所示。这4种结构在本书第1章已作了简要介绍,这里就局域网范畴作一些补充。总线型网中,各站

直接连在总线上，一般可以使用两种协议，一种是以太网使用的 CSMA/CD；另一种是令牌传递总线协议。环形网中最典型的是令牌环。星形网采用的是集中控制，由于集线器(Hub)和双绞线大量应用于局域网，使得星形网以及多级星形网获得了广泛使用。树形网只是总线型网的变形，可以说是总线型网的一种扩展。

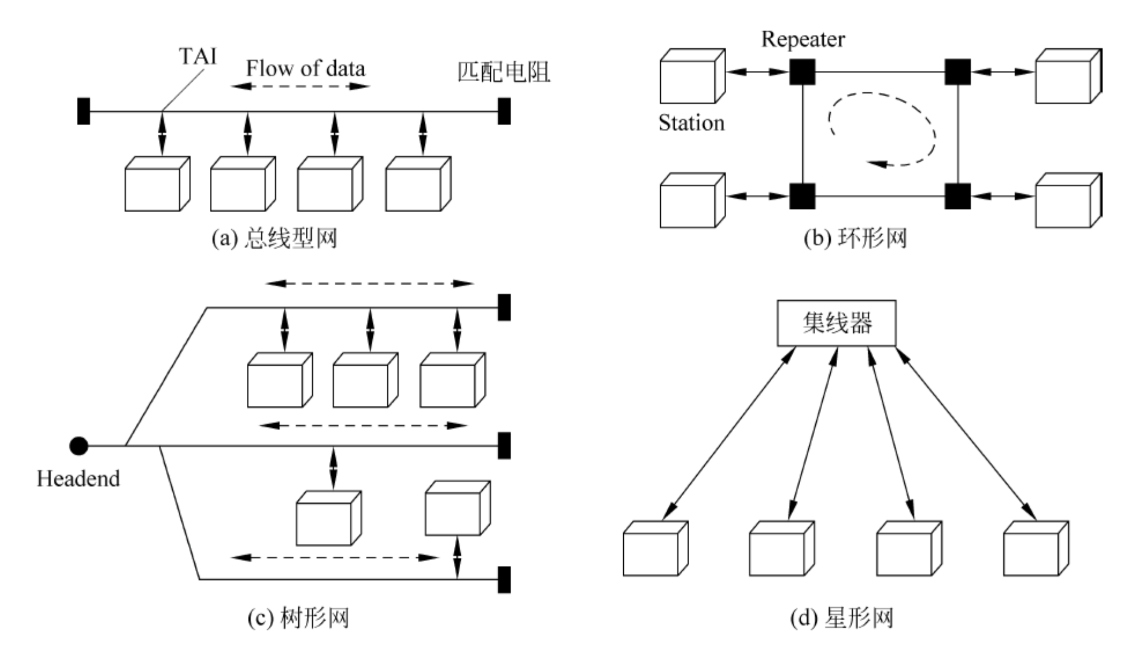


图 5.1 局域网的主要拓扑结构

拓扑的选择往往和传输介质的选择以及介质访问控制方法的确定紧密相关。选择拓扑时，应该考虑的主要因素有以下几点。

(1) 经济性。网络拓扑的选择直接决定了网络安装和维护的费用。不管选用什么样的传输介质，都需要进行安装。例如，安装电线沟、安装电线管道等。最理想的情况是建楼以前先进行安装，并考虑今后扩建的要求。安装费用的高低与拓扑结构的选择以及传输介质的选择、传输距离的确定有关。

(2) 灵活性。灵活性以及可扩充性也是选择网络拓扑结构时应充分重视的问题。任何一个网络，随着用户数的增加，网络应用的深入和扩大，网络新技术的不断涌现，特别是应用方式和要求的改变，网络经常需要加以调整。网络的可调整性与灵活性以及可扩充性都与网络拓扑直接相关。一般来说，总线型拓扑和环形拓扑要比星形拓扑的可扩充性好得多。

(3) 可靠性。网络的可靠性是任何一个网络的生命。网络拓扑决定了网络故障检测和故障隔离的方便性。

总之，选择局域网拓扑时，需要考虑的因素很多，这些因素同时影响网络的运行速度和网络软硬件接口的复杂程度等。

2. 局域网的传输介质

局域网可使用多种传输介质，包括双绞线、同轴电缆、光纤。在这里主要介绍有线传输介质。

1) 双绞线(TP)

将一对以上的双绞线封装在一个绝缘外套中，为了降低干扰，每对相互扭绕而成。分为非屏蔽双绞线(UTP)和屏蔽双绞线(STP)。局域网中 UTP 分为 3 类、4 类、5 类和超 5 类

4 种。

2）同轴电缆

由一根空心的外圆柱导体和一根位于中心轴线的内导线组成，两导体间用绝缘材料隔开。按直径分为粗缆和细缆。

(1) 粗缆：传输距离长，性能高但成本高，使用于大型局域网干线，连接时两端需终接器。

(2) 细缆：传输距离短，相对便宜，用 T 形头，与 BNC 网卡相连，两端安 50Ω 终端电阻。

按传输频带分为基带和宽带传输。

(1) 基带：数字信号，信号占整条信道，同一时间内能传送一种信号。

(2) 宽带：传送的是不同频率的信号。

3）光纤

应用光学原理，由光发送机产生光束，将电信号变为光信号，再把光信号导入光纤，在另一端由光接收机接收光纤上传来的光信号，并把它变为电信号，经解码后再处理。分为单模光纤和多模光纤。绝缘保密性好。

(1) 单模光纤：由激光作光源，仅有一条光通路，传输距离长，2km 以上。

(2) 多模光纤：由二极管发光，低速短距离，2km 以内。

5.1.4　局域网的工作模式

1. 专用服务器结构

专用服务器结构(Server-Based)又称为"工作站/文件服务器"结构，由若干台计算机工作站与一台或多台文件服务器通过通信线路连接起来组成工作站存取服务器文件，共享存储设备。文件服务器自然以共享磁盘文件为主要目的。

对于一般的数据传递来说已经够用了，但是当数据库系统和其他复杂而被不断增加的用户使用的应用系统到来时，服务器已经不能承担这样的任务了，因为随着用户的增多，为每个用户服务的程序也增多，每个程序都是独立运行的大文件，给用户感觉极慢，因此产生了客户机/服务器模式。

2. 客户机/服务器模式

客户机/服务器模式(Client/Server)，其中一台或几台较大的计算机集中进行共享数据库的管理和存取，称为服务器，而将其他的应用处理工作分散到网络中其他计算机上去做，构成分布式的处理系统，服务器控制管理数据的能力已由文件管理方式上升为数据库管理方式，因此，C/S 中的服务器也称为数据库服务器，注重于数据定义及存取安全后备与还原，并发控制及事务管理，执行诸如选择检索和索引排序等数据库管理功能，它有足够的能力做到把通过其处理后用户所需的那一部分数据而不是整个文件通过网络传送到客户机去，减轻了网络的传输负荷。C/S 结构是数据库技术的发展和普遍应用与局域网技术发展相结合的结果。

3. 对等式网络

在对等式网络(Peer-to-Peer)结构中，没有专用服务器，每一个工作站既可以起客户机作用，也可以起服务器作用。

5.2 局域网体系结构

在 20 世纪 80 年代初期，美国电气和电子工程师协会 IEEE 802 委员会首先制定出局域网的体系结构，即著名的 IEEE 802 参考模型，许多 IEEE 802 标准已成为 ISO 国际标准。

5.2.1 IEEE 802 参考模型

局域网的体系结构与 OSI 参考模型有相当大的区别，如图 5.2 所示，局域网只涉及 OSI 的物理层和数据链路层。为什么没有网络层及网络层以上的各层呢？首先，局域网是一种通信网，只涉及有关的通信功能，所以至多与 OSI 参考模型中的下 3 层有关。其次，由于局域网基本上采用共享信道的技术，所以也可以不设立单独的网络层。也就是说，不同局域网技术的区别主要在物理层和数据链路层，当这些不同的局域网需要在网络层实现互联时，可以借助其他已有的通用网络层协议（如 IP 协议）实现。

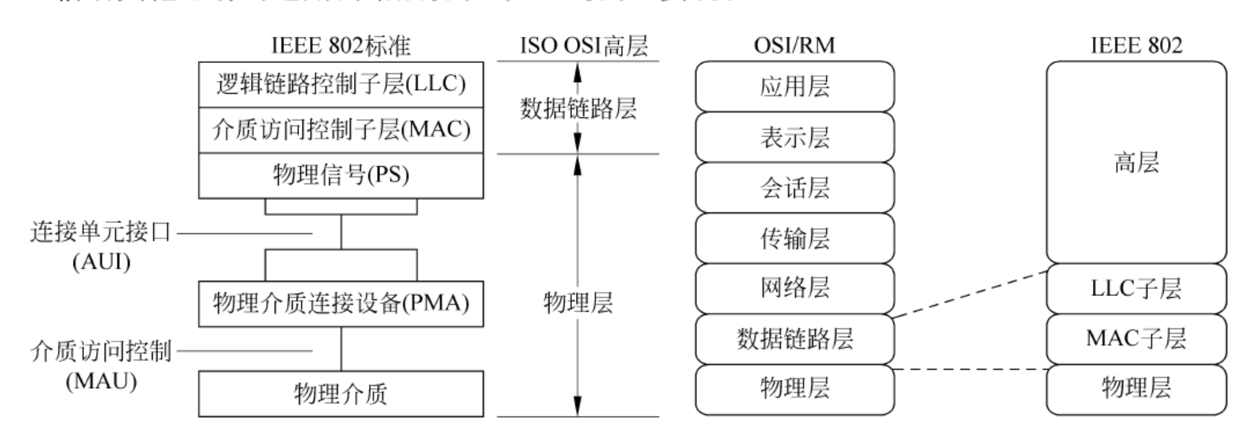


图 5.2 IEEE 802 局域网实现模型

1. 物理层

局域网的物理层是和 OSI 参考模型的物理层功能相当的，主要涉及局域网物理链路上原始比特流的传输，定义局域网物理层的机械、电气、规程和功能特性。如信号的传输与接收、同步序列的产生和删除等，物理连接的建立、维护、撤销等。

物理层还规定了局域网所使用的信号、编码、传输介质、拓扑结构和传输速率。例如，信号、编码可以采用曼彻斯特编码；传输介质可采用双绞线、同轴电缆、光纤甚至是无线传输介质；拓扑结构则支持总线型、星形、环形和混合型等，可提供多种不同的数据传输速率。物理层由以下 4 个部分组成。

（1）物理介质（PMD）：提供与线缆的物理连接。

（2）物理介质连接设备（PMA）：生成发送到线路上的信号，并接收线路上的信号。

（3）连接单元接口（AUI）。

（4）物理信号（PS）。

2. 数据链路层

局域网的数据链路层分为逻辑链路控制(Logical Link Control,LLC)和介质访问控制(Medium Access Control,MAC)两个功能子层。

局域网基本上采用共享介质环境,共享介质环境中的多个节点同时发送数据时会产生冲突,冲突(Collision)是指由于共享信道上同时有两个或两个以上的节点发送数据而导致信道上的信号波形不等于任何发送节点原始信号的情形。冲突会导致数据传输的失效,如图 5.3 所示,因而需要提供解决冲突的介质访问控制机制。

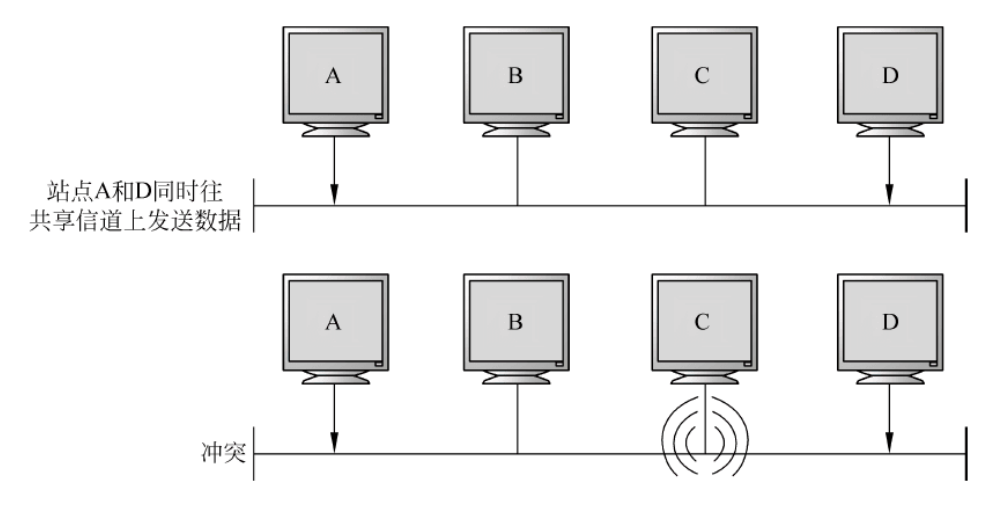


图 5.3 冲突与数据传输失效

但是,介质访问控制机制与物理介质、物理介质连接设备和物理拓扑等涉及物理实现的内容直接有关,也就是说,不同的局域网技术在介质访问控制上会有明显的差异。而这种差异是与计算机网络分层模型所要求的下层为上层提供服务、但必须屏蔽服务实现细节(即服务的透明性)是相违背的。为此,IEEE 802 标准的制定者考虑将局域网的数据链路层一分为二,成为 MAC 子层和 LLC 子层。

其中,MAC 子层负责介质访问控制机制的实现,即处理局域网中各站点对共享通信介质的争用问题,不同类型的局域网通常使用不同的介质访问控制协议,另外 MAC 子层还涉及局域网中的物理寻址;而 LLC 子层负责屏蔽掉 MAC 子层的不同实现,将其变成统一的 LLC 界面,从而向网络层提供一致的服务,LLC 子层向网络层提供的服务通过与网络层之间的逻辑接口实现,这些逻辑接口又被称为服务访问点(Service Access Point,SAP)。这样的局域网体系结构不仅使得 IEEE 802 标准更具有可扩充性,有利于其将来接纳新的介质访问控制方法和新的局域网技术,同时也不会使局域网技术的发展或变革影响到网络层。

尽管将局域网的数据链路层分成了 LLC 和 MAC 两个子层,但这两个子层是都要参与数据的封装和拆封过程的,而不是只由其中某一个子层来完成数据链路层帧的封装及拆封。在发送方,网络层下来的数据分组首先要加上 DSAP(Destination Service Access Point)和 SSAP(Source Service Access Point)等控制信息在 LLC 子层被封装成 LLC 帧,然后由 LLC 子层将其交给 MAC 子层,加上 MAC 子层相关的控制信息后被封装成 MAC 帧,最后由 MAC 子层交局域网的物理层完成物理传输;在接收方,则首先将物理的原始比特流还原成 MAC 帧,在 MAC 子层完成帧检测和拆封后变成 LLC 帧交给 LLC 子层,LLC 子层完成相应的帧检验和拆封工作将其还原成网络层的分组上交给网络层。从以上局域网的体系结构

可以看出，局域网数据链路层有两种不同的数据单元：LLC 帧和 MAC 帧。通常提到“帧”时是指 MAC 帧，而不是 LLC 帧。图 5.4 所示为 LLC 帧和 MAC 帧的关系示意图。

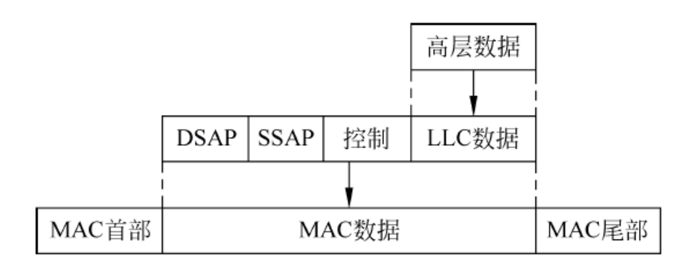


图 5.4 LLC 帧和 MAC 帧的关系

总之，局域网的 LLC 子层和 MAC 子层共同完成类似于 OSI 参考模型中的数据链路层功能，只是考虑到局域网的共享介质环境，在数据链路层的实现上增加了介质访问控制机制。

5.2.2 IEEE 802 系列标准

IEEE 在 1980 年 2 月成立了局域网标准化委员会（以下简称 IEEE 802 委员会），专门从事局域网的协议制定，形成了一系列的标准，称为 IEEE 802 标准。该标准已被国际标准化组织 ISO 采纳，作为局域网的国际标准系列，称为 ISO 802 标准。

IEEE 802 为局域网制定了一系列标准，主要有如下几种。

（1）IEEE 802.1：描述局域网体系结构以及寻址、网络管理和网络互联（1997）。

① IEEE 802.1G：远程 MAC 桥接（1998）。规定本地 MAC 网桥操作远程网桥的方法。

② IEEE 802.1H：在局域网中以太网 2.0 版 MAC 桥接（1997）。

③ IEEE 802.1Q：虚拟局域网（1998）。

（2）IEEE 802.2：定义逻辑链路控制（LLC）子层的功能与服务（1998）。

（3）IEEE 802.3：描述带冲突检测的载波监听多路访问（CSMA/CD）的访问方法和物理层规范（1998）。

① IEEE 802.3ab：描述 1000Base-T 访问控制方法和物理层技术规范（1999）。

② IEEE 802.3ac：描述 VLAN 的帧扩展（1998）。

③ IEEE 802.3ad：描述多重链接分段的聚合协议（Aggregation of Multiple Link Segments）（2000）。

④ IEEE 802.3i：描述 10Base-T 访问控制方法和物理层技术规范。

⑤ IEEE 802.3u：描述 100Base-T 访问控制方法和物理层技术规范。

⑥ IEEE 802.3z：描述 1000Base-X 访问控制方法和物理层技术规范。

⑦ IEEE 802.3ae：描述 10GBase-X 访问控制方法和物理层技术规范。

（4）IEEE 802.4：描述 Token Bus 访问控制方法和物理层技术规范。

（5）IEEE 802.5：描述 Token Ring 访问控制方法和物理层技术规范（1997）。

IEEE 802.5t：描述 100Mbps 高速标记环访问方法（2000）。

（6）IEEE 802.6：描述城域网（MAN）访问控制方法和物理层技术规范（1994）。1995 年又附加了 MAN 的 DQDB 子网上面向连接的服务协议。

(7) IEEE 802.7：描述宽带网访问控制方法和物理层技术规范。

(8) IEEE 802.8：描述 FDDI 访问控制方法和物理层技术规范。

(9) IEEE 802.9：描述综合语音、数据局域网技术(1996)。

(10) IEEE 802.10：描述局域网网络安全标准(1998)。

(11) IEEE 802.11：描述无线局域网访问控制方法和物理层技术规范(1999)。

(12) IEEE 802.12：描述 100VG-AnyLAN 访问控制方法和物理层技术规范。

(13) IEEE 802.14：描述利用 CATV 宽带通信的标准(1998)。

(14) IEEE 802.15：描述无线私人网(Wireless Personal Area Network，WPAN)。

(15) IEEE 802.16：描述宽带无线访问标准(Broadband Wireless Access Standards)，由两部分组成。

图 5.5 描述了 IEEE 802 标准的内部关系。

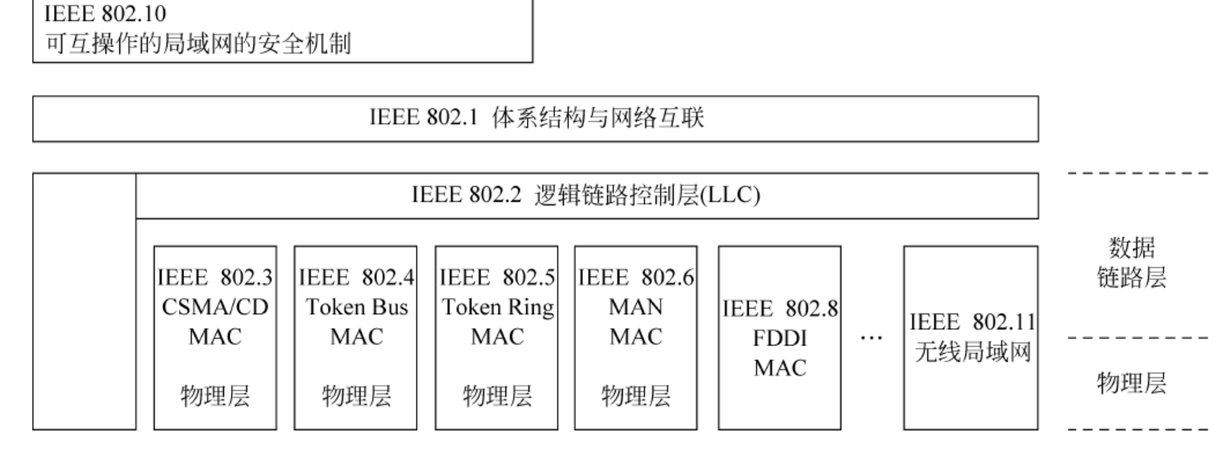


图 5.5　IEEE 802 标准的内部关系

从图 5.5 可以看出，IEEE 802 标准实际上是一个由一系列协议组成的标准体系。随着局域网技术的发展，该体系在不断地增加新的标准和协议，关于 IEEE 802.3 家族就随着以太网技术的发展出现了许多新的成员。

5.3　局域网中的介质访问控制

将传输介质的频带有效地分配给网上各站点用户的方法称为介质访问控制方法。介质访问控制方法是局域网最重要的一项基本技术，对局域网体系结构、工作过程和网络性能产生决定性影响。设计一个好的介质访问控制协议有 3 个基本目标：协议要简单，获得有效的通道利用率，公平合理地对待网上各站点的用户。介质访问控制方法主要是解决介质使用权的算法或机构问题，从而实现对网络传输信道的合理分配。

环形或总线型拓扑中，由于只有一条物理传输通道连接所有的设备，因此，连接到网络上的所有设备必须遵循一定的规则，才能确保传输媒体的正常访问和使用。常用的介质访问控制方法有具有冲突检测的载波监听多路访问/冲突检测 CSMA/CD(Carrier Sense Multiple Access/Collision Detection)、令牌环(Token Ring)及令牌总线(Token Bus)3 种技术。

5.3.1 信道分配问题

通常，可将信道分配方法划分为两类：静态分配方法和动态分配方法。

1. 静态分配方法

所谓静态分配方法，也是传统的分配方法，采用频分多路复用或时分多路复用的办法将单个信道划分后静态地分配给多个用户。

当用户站数较多或使用信道的站数在不断变化或者通信量的变化具有突发性时，频分多路复用方法的性能较差，因此，传统的静态分配方法，不完全适合计算机网络。

2. 动态分配方法

所谓动态分配方法，就是动态地为每个用户站点分配信道使用权。动态分配方法通常有3种：轮转、预约和争用。

(1) 轮转：使每个用户站点轮流获得发送的机会，这种技术称为轮转。它适合于交互式终端对主机的通信。

(2) 预约：预约是指将传输介质上的时间分隔成时间片，网上用户站点若要发送，必须事先预约能占用的时间片。这种技术适用于数据流的通信。

(3) 争用：若所有用户站点都能争用介质，这种技术称为争用。它实现起来简单，对轻负载或中等负载的系统比较有效，适合于突发式通信。争用方法属于随机访问技术，而轮转和预约的方法则属于控制访问技术。

5.3.2 载波监听多路访问/冲突检测

载波监听多路访问/冲突检测(CSMA/CD)是一种常用争用的方法来决定对媒体访问权的协议，这种争用协议只适用于逻辑上属于总线型拓扑结构的网络。在总线型网中，每个站点都能独立地决定帧的发送，若两个或多个站同时发送帧，就会产生冲突，导致所发送的帧都出错。因此，一个用户发送信息成功与否，在很大程度上取决于监测总线是否空闲的算法，以及当两个不同节点同时发送的分组发生冲突后所使用的中断传输的方法。总线争用技术可分为载波监听多路访问(CSMA)和具有冲突检测的载波监听多路访问(CSMA/CD)两大类。

1. 载波监听多路访问(CSMA)

载波监听多路访问(CSMA)的技术，也称为无听后说LBT(Listen Before Talk)。要传输数据的站点首先对介质媒体上有无载波进行监听，以确定是否有别的站点在传输数据。如果介质媒体空闲，该站点便可传输数据；否则，该站点将避让一段时间后再做尝试。这就需要有一种退避算法来决定避让的时间，常用的退避算法有非坚持、1-坚持、*P*-坚持3种。

1) 非坚持算法的算法规则

(1) 如果介质媒体是空闲的，则可以立即发送。

(2) 如果介质媒体是忙的，则等待一个由概率分布决定的随机重发延迟后，再重复步骤(1)。

采用随机的重发延迟时间可以减少冲突发生的可能性。非坚持算法的缺点是：即使有

几个站点都有数据要发送,但由于大家都在延迟等待过程中,致使介质媒体仍可能处于空闲状态,使用率降低。

2) 1-坚持算法的算法规则

(1) 如果介质媒体是空闲的,则可以立即发送。

(2) 如果介质媒体是忙的,则继续监听,直至检测到介质媒体是空闲,立即发送。

(3) 如果有冲突(在一段时间内未收到肯定的回复),则等待一随机量的时间,重复步骤(1)~(2)。

这种算法的优点是:只要介质媒体空闲,站点就立即可发送,避免了介质媒体利用率的损失;其缺点是:假若有两个或两个以上的站点有数据要发送,冲突就不可避免。

3) P-坚持算法的算法规则

(1) 监听总线,如果介质媒体是空闲的,则以 P 的概率发送,而以$(1-P)$的概率延迟一个时间单位。一个时间单位通常等于最大传播时延的 2 倍。

(2) 延迟一个时间单位后,再重复步骤(1)。

(3) 如果介质媒体是忙的,继续监听直至介质媒体空闲并重复步骤(1)。

P-坚持算法是一种既能像非坚持算法那样减少冲突,又能像 1-坚持算法那样减少介质媒体空闲时间的折中方案。问题在于如何选择 P 的值,这要考虑到避免重负载下系统处于的不稳定状态。假如介质媒体是忙时,有 N 个站数据等待发送,一旦当前的发送完成时,将要试图传输的站的总期望数为 NP。如果选择 P 过大,使 $NP>1$,表明有多个站点试图发送,冲突就不可避免。最坏的情况是,随着冲突概率的不断增大,而使吞吐量降低到零。所以必须选择适当 P 值使 $NP<1$。当然 P 值选得过小,则介质媒体利用率又会大大降低。

2. 具有冲突检测的载波监听多路访问(CSMA/CD)

在 CSMA 中,由于信道传播时延的存在,即使总线上两个站点没有监听到载波信号而发送帧时,仍可能会发生冲突。由于 CSMA 算法没有冲突检测功能,即使冲突已发生,仍然将已破坏的帧发送完,使总线的利用率降低。

一种 CSMA 的改进方案是使发送站点传输过程中仍继续监听媒体,以检测是否存在冲突。如果发生冲突,信道上可以检测到超过发送站点本身发送的载波信号的幅度,由此判断出冲突的存在。一旦检测到冲突,就立即停止发送,并向总线上发一串拥塞信号,用以通知总线上其他各有关站点。这样,通道容量就不致因白白传送已受损的帧而浪费,可以提高总线的利用率。这种方案称为载波监听多路访问/冲突检测协议,简写为 CSMA/CD,“载波监听”指以太网的网络接口卡监听网络,直到没有节点正在发送时,才开始发送数据;“多路访问”指多个节点连接到同一个网络上,并能同时检测信道,线路空闲时任何节点都能发送数据。这种协议已被广泛应用于局域网中。

CSMA/CD 的代价是用于检测冲突所花费的时间。对于基带总线而言,最坏情况下用于检测一个冲突的时间等于任意两个站之间传播时延的两倍。从一个站点开始发送数据到另一个站点开始接收数据,也即载波信号从一端传播到另一端所需的时间,称为信号传播时延。

$$\text{信号传播时延}(\mu s)=\text{两站点的距离}(m)\div\text{信号传播速率}(200m/\mu s)$$

假定 A、B 两个站点位于总线两端,两站点之间的最大传播时延为 tp。当 A 站点发送数据后,经过接近于最大传播时延 tp 时,B 站点正好也发送数据,此时冲突便发生。发生冲

突后，B站点立即可检测到该冲突，而A站点需再经过一份最大传播时延tp后，才能检测出冲突。也即最坏情况下，对于基带CSMA/CD来说，检测出一个冲突的时间等于任意两个站之间最大传播时延的两倍(2tp)。

数据帧从一个站点开始发送，到该数据帧发送完毕所需的时间和为数据传输时延；同理，数据传输时延也表示一个接收站点开始接收数据帧，到该数据帧接收完毕所需的时间。

数据传输时延(s)＝数据帧长度(b)÷数据传输速率(bps)

若不考虑中继器引入的延迟，数据帧从一个站点开始发送，到该数据帧被另一个站点全部接收所需的总时间，等于数据传输时延与信号传播时延之和。

从以上分析可知，为了确保发送数据站点在传输时能检测到可能存在的冲突，数据帧的传输时延至少要两倍于传播时延。换句话说，要求分组的长度不短于某个值，否则在检测出冲突之前传输已经结束，但实际上分组已被冲突所破坏。

冲突检测设计中注意两个问题：帧的大小和传输距离。

若帧过大，站点将独占介质。另一方面，冲突检测要求帧不能太小，以确保站点在一个帧发完之前能够检测到一次冲突。如果在发送完毕之后才检测冲突，它就无法判断该冲突是否与它发送出去的帧有关。帧可能已经到达了目的地，而另外两个帧冲突了。

一个帧可以多小？答案取决于检测到一次冲突所需要的时间。有时冲突几乎立即被检测得到；有时一个信号可能会经过很长的距离才撞上另一个信号。因此，在最坏情况下，检测到一次冲突所需要的时间等于信号在介质中经过最长距离所需要时间的两倍。

例：如下假设。

(1) 站点在一条速率为10Mbps的同轴电缆上发送帧。

(2) 电缆上两个站点的最远距离是2km。

(3) 信号在电缆上的传播速率为200m/μs。

最坏情况下，一个帧与另一个帧相撞之前传播了2km(耗时10μs)，接着被破坏的帧再走2km回到发送站点，来回共耗时20μs，所以发送一个帧的最少时间是20μs。10Mbps的数据传输速率相当于每微秒发送10个比特，20μs就是200b。这意味着一个帧至少应该有200b，即200/8＝25(B)。

在CSMA/CD算法中，一旦检测到冲突并发完阻塞信号后，为了降低再次冲突的概率，需要等待一个随机时间，然后再使用CSMA方法试图传输。为了保证这种退避操作维持稳定，采用了一种称为二进制指数退避算法。规则如下：

(1) 对每个数据帧，当第一次发生冲突时，设置一个参量$L=2$；

(2) 退避间隔取1～L个时间片中的一个随机数，1个小时片等于两站之间的最大传播时延的两倍；

(3) 当数据帧再次发生冲突，则将参量L加倍；

(4) 设置一个最大重传次数，超过该次数，则不再重传，并报告出错。

如果两台或多台计算机在冲突后恰好选择几乎相同的延迟，那么它们几乎同时开始传输，导致第二次冲突。为了防止一连串的冲突，以太网要求每台计算机在每次冲突后把选择延迟的范围加倍。这样，计算机在第一次冲突后从0～D选择一个随机延迟，在第二次冲突后从0～$2D$选择一个随机即延迟，在第三次后从0～$4D$选择，以此类推。在几次冲突之后，选择随机值的范围变得很大，没有冲突传输的概率就变得很大。

二进制指数退避算法是按后进先出 LIFO(List In First Out)的次序控制的,即未发生冲突或很少发生冲突的数据帧,具有优先发送的概率;而发生过多次冲突的数据帧,发送成功的概率就更少。IEEE 802.3 就是采用二进制指数退避算法和 1-坚持算法的 CSMA/CD 媒体访问控制方法。这种方法在低负荷时,如媒体空闲时,要发送数据帧的站点能立即发送;在重负荷时,仍能保证系统的稳定性。由于在媒体上传播的信号会衰减,为确保能检测出冲突信号,CSMA/CD 总线型网限制一段无分支电缆的最大长度为 500m。

总之,CSMA/CD 采用的是一种"有空就发"的竞争型访问策略,因而不可避免地会出现信道空闲时多个站点同时争发的现象,无法完全消除冲突,只能是采取一些措施减少冲突,并对产生的冲突进行处理。因此采用这种协议的局域网环境不适合对实时性要求较强的网络应用。

5.3.3　令牌环介质访问控制

Token Ring 是令牌传输环(Token Passing Ring)的简写。令牌环介质访问控制方法是通过在环形网上传输令牌的方式来实现对介质的访问控制。只有当令牌传输至环中某站点时,它才能利用环路发送或接收信息。当环线上各站点都没有帧发送时,令牌标记为 01111111,称为空标记。当一个站点要发送帧时,需等待令牌通过,并将空标记置换为忙标记 01111110,紧跟着令牌,用户站点把数据帧发送至环上。由于是忙标记,所以其他站点不能发送帧,必须等待。

发送出去的帧将随令牌沿环路传输下去。在循环一周又回到原发送站点时,由发送站点将该帧从环上移去,同时将忙标记换为空标记,令牌传至后面站点,使之获得发送的许可权。发送站点在从环中移去数据帧的同时还要检查接收站载入该帧的应答信息,若为肯定应答,说明发送的帧已被正确接收,完成发送任务。若为否定应答,说明对方未能正确收到所发送的帧,原发送站点需在带空标记的令牌第二次到来时,重发此帧。采用发送站从环上收回帧的策略,不仅具有对发送站点自动应答的功能,而且还具有广播特性,即可有多个站点接收同一数据帧。

接收帧的过程与发送帧不同,当令牌及数据帧通过环上站点时,该站将帧携带的目标地址与本站地址相比较。若地址符合,则将该帧复制下来放入接收缓冲器中,待接收站正确接收后,即在该帧上载入肯定应答信号;若不能正确接收,则载入否定应答信号,之后再将该帧送入环上,让其继续向下传输。若地址不符合,则简单地将数据帧重新送入环中。所以当令牌经过某站点而它既不发送信息,又无处接收时,会稍经延迟,继续向前传输。

在系统负载较轻时,由于站点需等待令牌到达才能发送或接收数据,因此效率不高。但若系统负载较重,则各站点可公平共享介质,效率较高。为避免所传输数据与标记形式相同而造成混淆,可采用前面所讲过的位填入技术,以区别数据和标记。

使用令牌环介质访问控制方法的网络,需要有维护数据帧和令牌的功能。如可能会出现因数据帧未被正确移去而始终在环上传输的情况。也可能出现令牌丢失或只允许一个令牌的网络中出现了多个令牌等异常情况。解决这类问题的办法是在环中设置监控器,对异常情况进行检测并消除。令牌环网上的各个站点可以设置成不同的优先级,允许具有较高优先权的站申请获得下一个令牌权。

归纳起来,在令牌环中主要有下面 3 种操作。

(1) 截获令牌并且发送数据帧。如果没有节点需要发送数据,令牌就由各个节点沿固定的顺序逐个传递;如果某个节点需要发送数据,它要等待令牌的到来,当空闲令牌传到这个节点时,该节点修改令牌帧中的标志,使其变为"忙"的状态,然后去掉令牌的尾部,加上数据,成为数据帧,发送到下一个节点。

(2) 接收与转发数据。数据帧每经过一个节点,该节点就比较数据帧中的目的地址,如果不属于本节点,则转发出去;如果属于本节点,则复制到本节点的计算机中,同时在帧中设置已经复制的标志,然后向下一个节点转发。

(3) 取消数据帧并且重发令牌。由于环形网在物理上是个闭环,一个帧可能在环中不停地流动,所以必须清除。当数据帧通过闭环重新传到发送节点时,发送节点不再转发,而是检查发送是否成功。如果发现数据帧没有被复制(传输失败),则重发该数据帧;如果发现传输成功,则清除该数据帧,并且产生一个新的空闲令牌发送到环上。

令牌环介质访问控制,如图 5.6 所示。令牌环主要优点是它提供对传输介质访问的灵活控制。而且在负载很重的情况下,这种令牌环的控制策略是高效和公平的。它的主要缺点一个是在轻负载的情况下,由于传输信包前必须等待一个空令牌的到来,这样造成了一些低效率;另一个是需要对令牌进行维护,一旦令牌丢失,环形网便不能再运行,所以在环路上要设置一个站点作为环上的监控站点,来保证环上有且仅有一个令牌。

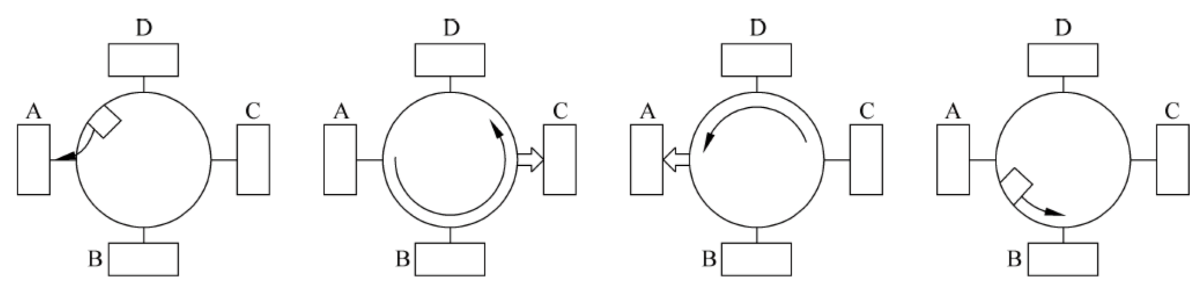


图 5.6 令牌环介质访问控制

5.3.4 令牌总线

CSMA/CD 访问控制采用总线争用方式,具有结构简单、在轻负载下延迟小等优点,但随着负载的增加,冲突概率增加,性能将明显下降。采用令牌环介质访问控制具有重负载下利用率高、网络性能对距离不敏感以及具有公平访问等优越性能,但环形网结构复杂,存在检错和可靠性等问题。令牌总线介质访问控制是在综合了以上两种介质访问控制优点的基础上形成的一种访问控制方法,IEEE 802.4 提出的就是令牌总线介质访问控制方法的标准。令牌总线使所有的站接入一个线性电缆,并在逻辑上组织为一个环,如图 5.7 所示。当逻辑环初始化时,最高地址的站可以发送第一个帧,以后它传一个称为令牌的特殊的帧给它的邻居,允许它去发送。令牌沿环环行,只有令牌的持有者允许发送帧,不会有拥塞发生。

这种方式和 CSMA/CD 方式一样,采用总线型网络拓扑,但不同的是网上各工作站按一定顺序形成一个逻辑环。每个工作站在环中均有一个指定逻辑位置,末站的后站就是首

站（即首尾相连），每站都了解其先行站和后继站的地址，总线上各站的物理位置与逻辑位置无关，如图 5.8 所示。

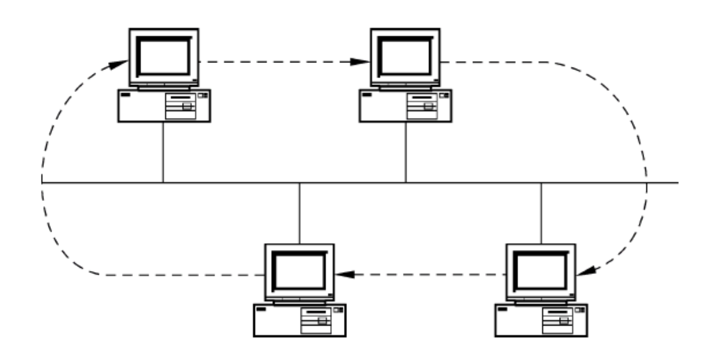


图 5.7 IEEE 802.4 令牌总线

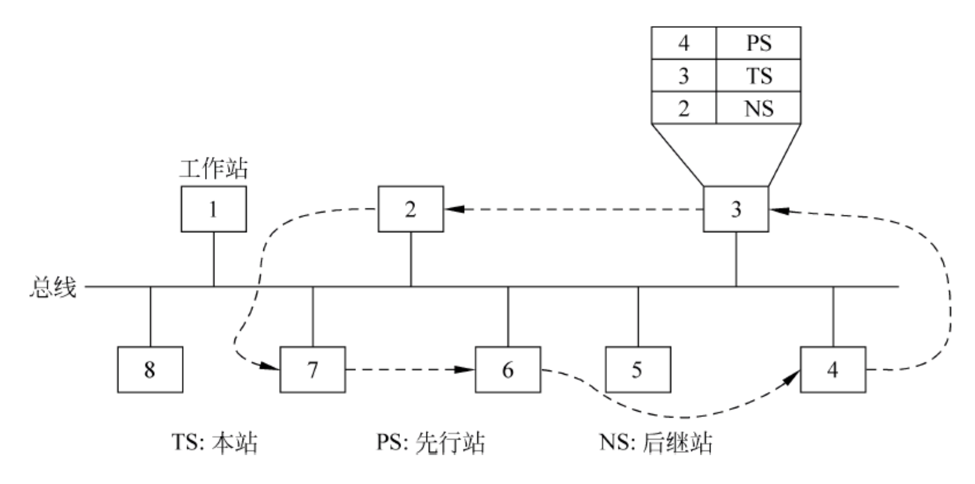


图 5.8 令牌总线工作示意图

从物理上看，所有站点都挂到一根总线上。但在逻辑上，所有工作的站点都被组织在一个逻辑环内。和令牌环不同之处是，总线型网中的令牌需要携带地址，并且操作比较复杂。

令牌传送顺序仅取决于各站在逻辑环上的位置，而与其在总线上的物理位置无关。因为总线型网是广播型的，每站都可收到总线上传输的帧，但它们只接收本站为目的站的帧。当令牌经过本站而无数据发送时，则本站一收到令牌就立即将其传递给下一站。

令牌总线型网为不同负载提供不同的优先级别服务，当一个站点得到令牌后，按优先级别发送数据，而且优先级别越高，分配发送的时间越长。

站点连到电缆上的实际顺序并不重要，因为电缆是固有的广播介质，所以每个站点都可以收到所有的帧，但把不是发给它的帧丢掉。当站点传递令牌时，只要向环上的逻辑邻居发送一个令牌帧，无须考虑该邻居实际上位于电缆何处。

称为令牌的控制帧调整访问的权利。令牌帧包含目的地址，收到令牌的工作站在一段规定的时间内被授予对介质的控制权，因而该站可以发送一帧或多帧信息。

与令牌环形网一样，在令牌总线型网中，只有令牌持有者才能控制总线，才有信息发送权，它可以发送一帧或多帧信息。当该站发送完成或时间到，它就将令牌传给逻辑环中的下一个工作站，由于环上只有一个令牌依次传递，故不会出现碰撞。网上的工作站也可以不进入逻辑环而成为非活动节点，这些工作站仅能响应询问或请求回答。

归纳起来，在令牌总线中主要有下面几种管理操作。

（1）环初始化，即生成一个顺序访问的次序。网络开始启动时，或由于某种原因，在运行中所有站点不活动的时间超过规定的时间，都需要进行逻辑环的初始化。初始化的过程是一个争用的过程，争用结果只有一个站能取得令牌，其他的站点用站插入的算法插入。

（2）令牌传递算法。逻辑环按递减的站地址次序组成，刚发完帧的站点将令牌传递给后继站，后继站应立即发送数据或令牌帧，原先释放令牌的站监听到总线上的信号，便可确认后继站已获得令牌。

（3）站插入环算法。必须周期性地给未加入环的站点以机会，将它们插入逻辑环的适当位置中。如果同时有几个站要插入，可采用带有响应窗口的争用处理算法。

（4）站退出环算法。可以通过将其前趋站和后继站连到一起的办法，使不活动的站退出逻辑环，并修正逻辑环递减的站地址次序。

（5）故障处理。网络可能出现错误，这包括令牌丢失引起断环、重复地址、产生多个令牌等。网络需对这些故障作出相应的处理。

令牌总线介质访问控制方法是能够保证每个工作站在某一定时间间隔内访问介质；可以用数种方法建立优先权；送去时间是确定的，适用于实时性较强的场合，但是实现起来比CSMA/CD复杂。

5.4 以 太 网

以太网（Ethernet）是一种产生较早且使用相当广泛的局域网，由美国Xerox（施乐）公司于20世纪70年代初期开始研究并于1975年推出。由于它具有结构简单、工作可靠、易于扩展等优点，因而得到了广泛的应用。1980年，美国Xerox、DEC与Intel三家公司联合提出了以太网规范，这是世界上第一个局域网的技术标准。后来的以太网国际标准IEEE 802.3就是参照以太网的技术标准建立的，两者基本兼容。为了与后来的快速以太网相区别，通常又将这种按IEEE 802.3规范生产的以太网产品称为传统以太网，简称为以太网。

5.4.1 以太网特征及分类

1. 以太网特征

以太网具有的主要技术特征如下。

（1）以太网是基带网，它采用基带传输技术。

（2）以太网的标准是IEEE 802.3，它使用CSMA/CD访问方法。

（3）以太网是一种共享型网络，网络上的所有站点共享传输媒体和带宽。当利用率到达40%时，网络的响应速度明显降低。

（4）以太网是广播式网络，因此，它具有广播式网络的全部特点。

（5）以太网的数字信号采用曼彻斯特编码方案，快速以太网采用4B/5B编码方案。

（6）以太网支持传输介质类型有50Ω基带同轴电缆、无屏蔽双绞线和光纤。

(7) 以太网所构成的拓扑结构主要是总线型和星形。

(8) 有多种以太网标准，传输速率为 10Mbps、100Mbps、1000Mbps、10 000Mbps 以及更高。

(9) 以太网是可变长帧，长度为 64～1514B。

(10) 以太网技术先进，价格低廉，易扩展、易维护、易管理。

2. 以太网分类

1) 标准以太网

最初以太网只有 10Mbps 的吞吐量，它所使用的是 CSMA/CD 的访问控制方法，通常把这种最早期的 10Mbps 以太网称为标准以太网。以太网主要有两种传输介质，那就是双绞线和同轴电缆。所有的以太网都遵循 IEEE 802.3 标准，下面列出的是 IEEE 802.3 的一些以太网标准，在这些标准中前面的数字表示传输速率，单位是 Mbps，最后的一个数字表示单段网线长度(基准单位是 100m)，Base 表示“基带”的意思，Broad 代表“带宽”。

(1) 10Base-5：使用粗同轴电缆，最大网段长度为 500m，基带传输方法。

(2) 10Base-2：使用细同轴电缆，最大网段长度为 185m，基带传输方法。

(3) 10Base-T：使用双绞线电缆，最大网段长度为 100m。

(4) 1Base-5：使用双绞线电缆，最大网段长度为 500m，传输速率为 1Mbps。

(5) 10Broad-36：使用同轴电缆(RG-59/UCATV)，最大网段长度为 3600m，是一种宽带传输方式。

(6) 10Base-F：使用光纤传输介质，传输速率为 10Mbps。

2) 快速以太网

随着网络的发展，传统标准的以太网技术，已难以满足日益增长的网络数据流量速度需求。在 1993 年 10 月以前，对于要求 10Mbps 以上数据流量的 LAN 应用，只有光纤分布式数据接口(FDDI)可供选择，但它是一种价格非常昂贵的、基于 100Mbps 光缆的 LAN。1993 年 10 月，Grand Junction 公司推出了世界上第一台快速以太网(Fast Ethernet)集线器 Fastch10/100 和网络接口卡 FastNIC100，快速以太网技术正式得以应用。随后 Intel、SynOptics、3COM、BayNetworks 等公司也相继推出自己的快速以太网装置。与此同时，IEEE 802 工程组也对 100Mbps 以太网的各种标准，如 100Base-TX、100Base-T4、MII、中继器、全双工等标准进行了研究。1995 年 3 月 IEEE 宣布了 IEEE 802.3u 100Base-T 快速以太网标准，就这样开始了快速以太网的时代。

快速以太网与原来在 100Mbps 带宽下工作的 FDDI 相比它具有许多的优点，最主要体现在快速以太网技术可以有效地保障用户在布线基础实施上的投资，它支持 3 类、4 类、5 类双绞线以及光纤的连接，能有效地利用现有的设施。

快速以太网的不足其实也是以太网技术的不足，那就是快速以太网仍是基于载波监听多路访问/冲突检测(CSMA/CD)技术，当网络负载较重时，会造成效率的降低，当然这可以使用交换技术来弥补。

100Mbps 快速以太网标准又分为 100Base-TX 、100Base-FX、100Base-T4 三个子类。

(1) 100Base-TX：是一种使用 5 类数据级无屏蔽双绞线或屏蔽双绞线的快速以太网技术。它使用两对双绞线，一对用于发送数据，一对用于接收数据。在传输中使用 4B/5B 编码方式，信号频率为 125MHz。符合 EIA586 的 5 类布线标准和 IBM 的 SPT 1 类布线标

准。使用同10Base-T相同的RJ-45连接器。它的最大网段长度为100m。它支持全双工的数据传输。

(2) 100Base-FX：是一种使用光缆的快速以太网技术，可使用单模光纤和多模光纤(62.5和125um)，多模光纤连接的最大距离为550m。单模光纤连接的最大距离为3000m。在传输中使用4B/5B编码方式，信号频率为125MHz。它使用MIC/FDDI连接器、ST连接器或SC连接器。它的最大网段长度为150m、412m、2000m或更长至10km，这与所使用的光纤类型和工作模式有关，它支持全双工的数据传输。100Base-FX特别适用于有电气干扰的环境、较大距离连接，或高保密环境等情况下。

(3) 100Base-T4：是一种可使用3类、4类、5类无屏蔽双绞线或屏蔽双绞线的快速以太网技术。它使用4对双绞线，3对用于传送数据，1对用于检测冲突信号。在传输中使用8B/6T编码方式，信号频率为25MHz，符合EIA586结构化布线标准。它使用与10Base-T相同的RJ-45连接器，最大网段长度为100m。

3) 千兆以太网

千兆以太网技术作为最新的高速以太网技术，给用户带来了提高核心网络的有效解决方案，这种解决方案的最大优点是继承了传统以太网技术价格便宜的优点。

千兆以太网技术仍然是以太技术，它采用了与10M以太网相同的帧格式、帧结构、网络协议、全/半双工工作方式、流控模式以及布线系统。由于该技术不改变传统以太网的桌面应用、操作系统，因此可与10M或100M的以太网很好地配合工作。升级到千兆以太网不必改变网络应用程序、网管部件和网络操作系统，能够最大限度地投资保护，因此该技术的市场前景十分看好。

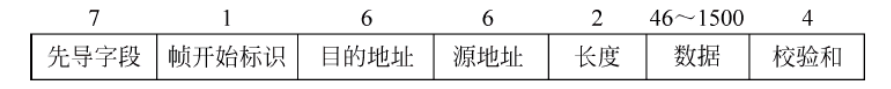
千兆以太网技术有两个标准：IEEE 802.3z和IEEE 802.3ab。IEEE 802.3z制定了光纤和短程铜线连接方案的标准，目前已完成了标准制定工作。IEEE 802.3ab制定了5类双绞线上较长距离连接方案的标准。

(1) IEEE 802.3z。IEEE 802.3z工作组负责制定光纤(单模或多模)和同轴电缆的全双工链路标准。IEEE 802.3z定义了基于光纤和短距离铜缆的1000Base-X，采用8B/10B编码技术，信道传输速率为1.25Gbps，去耦后实现1000Mbps传输速率。

(2) IEEE 802.3ab。IEEE 802.3ab工作组负责制定基于UTP的半双工链路的千兆以太网标准，产生IEEE 802.3ab标准及协议。IEEE 802.3ab定义基于5类UTP的1000Base-T标准，其目的是在5类UTP上以1000Mbps传输速率传输100m。

3. 以太网的帧格式与接收过程

图5.9给出了IEEE 802.3帧格式。



7	1	6	6	2	46～1500	4
先导字段	帧开始标识	目的地址	源地址	长度	数据	校验和

图5.9 IEEE 802.3帧格式

其中有关字段的说明如下。

(1) 先导字段：长度为7B，每个字节的内容为10101010，用于接收方与发送方的时钟同步。

(2) 帧开始标识：长度为1B，内容为10101011，标志帧的开始。

(3) 目的地址和源地址：均为6B。分别表示接收节点和目标节点的MAC地址。当目的地址为二进制全1(相当于12b的十六进制F)时，表示该帧要被传送至网络上的所有节点，即所谓的“广播”。

(4) 长度：长度为2B，用于指明数据字段中的字节数，取值范围为0～1500。IEEE 802.3中数据长度可为0，当数据长度小于46B时，需要使用填充字段以达到帧长度≥64B的要求。

(5) 数据：长度为46～1500B，在CSMA/CD中，当收发器检测到冲突时，就将当前帧的其余部分丢弃，这样残余帧就会一直出现在信道上。为了区分有效帧和残余帧，IEEE 802.3规定有效帧中从目的地址到校验和字段的最短长度为64B。如果帧的数据部分少于46B，必须使用填充字段以达到所要求的最短长度。

在以太网中，只要一个节点成功地利用总线发送了一个帧，那么其他节点都处于接收状态，节点接收以太网帧的流程如图5.10所示。

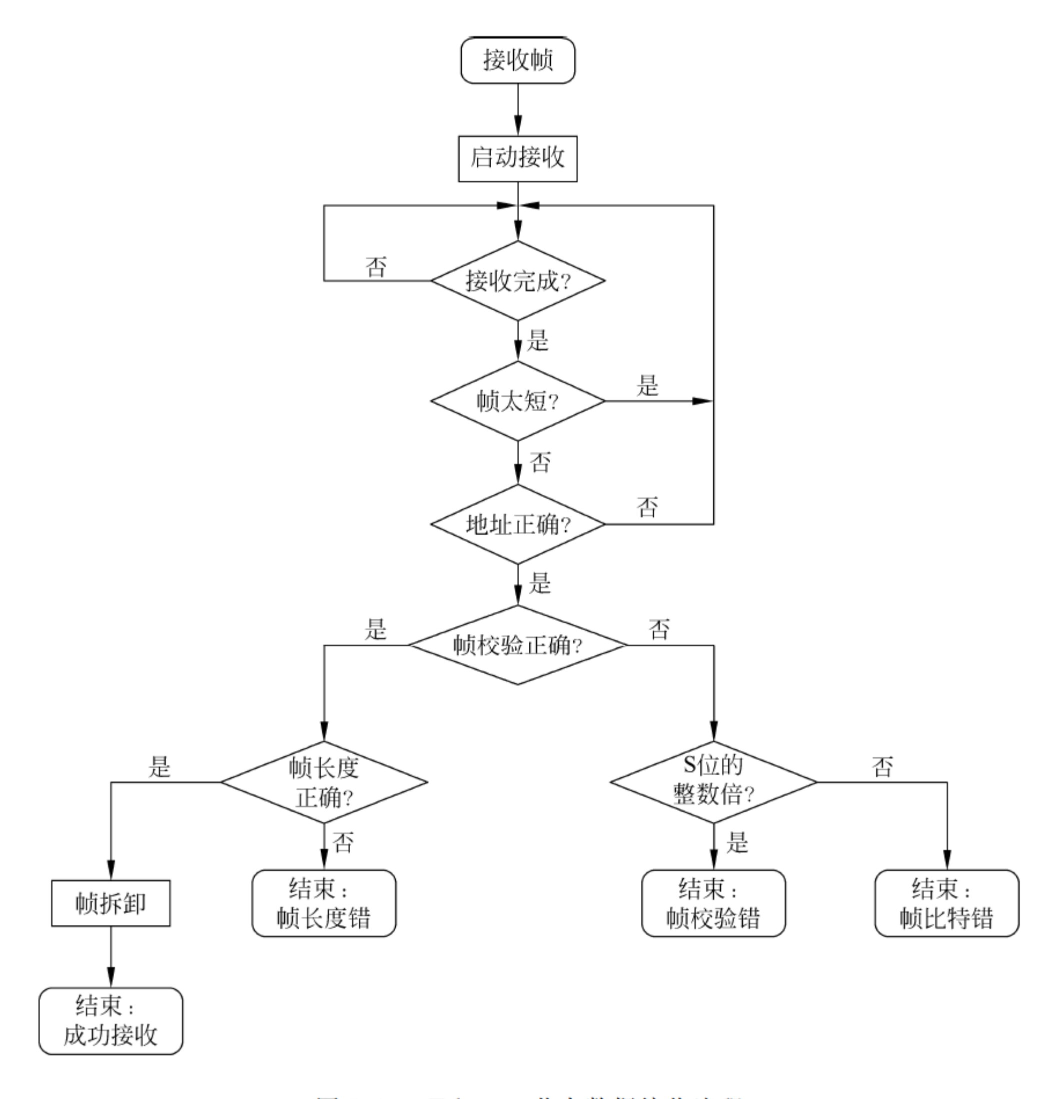


图5.10 Ethernet节点数据接收流程

首先，接收节点要判断接收帧的长度。如果所接收到的帧长小于以太网所定义的最小帧长度，则表明有冲突发生，应丢弃该帧，并重新等待接收；如果所接收到的帧长大于所规

定的最小帧长度，则检查帧中的目的地址是否与本节点的 MAC 地址匹配。

如果目的地址为单播地址，且目的地址为本节点的地址，则匹配；如果目的地址为组播地址，且本节点地址是组播组中的一员，则匹配；如果目的地址为广播地址，那么也匹配。注意，只有目的地址与本节点地址匹配的帧才被节点所接收。

节点接收帧之后，还要进行 CRC 校验。如果 CRC 校验正确，则进一步检验 LLC 数据长度是否正确。如果 CRC 正确但 LLC 数据长度不正确，则报告“帧长度错”，如果 CRC 校验与 LLC 数据长度都正确，则将帧中 LLC 数据送 LLC 子层，报告“成功接收”，进入结束状态。

如果 CRC 校验中发现错误，则首先判断接收帧是不是 8b 的整数倍；如果帧的长度是 8b 的整数倍，则表示传输过程中没有发现比特丢失或对错位，此时应记录“帧校验错”并进入结束状态。如果帧长度不是 8b 的整数倍，则报告“帧比特位错”并进入结束状态。

5.4.2 以太网组网技术

IEEE 802.3 委员会在定义可选的物理配置方面表现了极大的多样性和灵活性。为了区分各种可选用的实现方案，该委员会给出了一种简明的表示方法。

＜数据传输速率(Mbps)＞ ＜信号方式＞ ＜最大段长度(百米)＞

如 10Base-5、10Base-2、10Broad-36。10Base-5 标准表示传输速率是 10Mbps，使用基带信号传送数据，最大距离是 500m。10Base-2 使用细缆总线和 BNC-T 连接器。但 10Base-T 有些例外，其中的 T 表示双绞线、光纤。

在以太网的发展过程中，曾出现 10Base-5、10Base-2、10Base-T、100Base-T、100Base-F 等类型，下面简要介绍。

1. 10Base-5(粗缆以太网)

网卡通过 DB-15 型连接器与收发器电缆(又称 AUI 电缆)相连，然后再连接到收发器上，收发器再和粗缆连接。这里 10 表示信号在电缆上的传输速率为 10Mbps，Base 表示电缆上的信号是基带信号，5 表示每一段电缆的最大长度为 500m。由于采用的传输媒体是特性阻抗为 50Ω、直径为 10mm 的同轴电缆，所以这种以太网通常称为粗缆以太网。

同 10Base-2 技术比较起来，二者的主要区别在于：10Base-5 采用 10mm 的粗缆作为传输介质，其单个网段最大长度可以达到 500m；而 10Base-2 的传输介质用的是 5mm 的细缆，其单个网段的最大距离只能达到 185m 左右。另外，由于采用了不同的传输介质，所以两个网络使用的介质连接器也不相同。

10Base-5 的技术规范如下。

(1) 每个网段的最大距离为 500m。

(2) 在每个网段可以使用中继器扩充网络覆盖范围。最多使用 4 个中继器，连接 5 个线段，使整个 10Base-5 网络的最大长度达到 2500m。

(3) 在每个线段中连接的工作站的数目最多为 100 个，其中中继器也算为节点数目。

(4) 两个收发器之间的最短距离为 2.5m，收发器电缆最长为 50m。

10Base-5 网络由于采用粗缆作为传输介质，并且在每一个节点处需要收发器，因此一般很少在局域网中采用纯粹的 10Base-5 组成网络，但是由于其传输距离较远，因此有时可

以采用混合连接的方法,用粗缆来延长网络的覆盖范围。

2. 10Base-2(细缆以太网)

10Base-2 网络结构示意图如图 5.11 所示。用 BNC T 形连接器(它有 3 个头,外观像 T,以下简称 T 形头),其两个反向头连接两段电缆,中间的一个头连接到网卡外露的 BNC 插座上。它采用 50Ω、直径为 5mm 的细同轴电缆,通常称为细缆以太网。2 表示每个网段长度约为 200m,准确的为 185m。

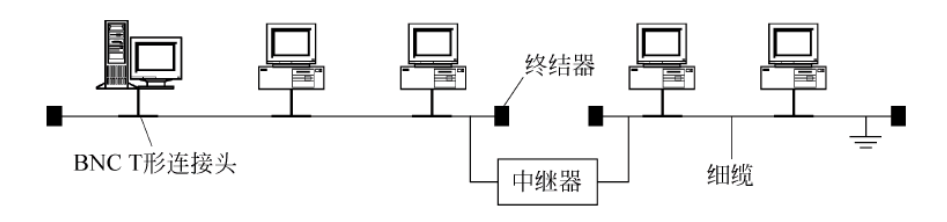


图 5.11 10Base-2 网络结构示意图

10Base-2 也称为细缆以太网。在 10Base-2 被开发出来之前,以太网仅使用粗同轴电缆。细缆同粗缆相比,其灵活性更高且更易于安装。在使用这种网络时,必须遵守以下的技术规范。

(1) 一个 10Base-2 以太网段的最大长度是 185m(607 英尺)。10Base-2 网络使用两个 50Ω 电阻器终止网络的两端以防止信号反射。

(2) 每个网络段最多包含 30 个节点,其中每个节点通过一个 BNC 连接器连接到总线上(参见图 5.11)。

(3) BNC T 形连接器附着在 10Base-2 网络专用网络接口卡的 BNC 连接器上。每个 BNC T 形连接器之间的距离应至少是 0.5m。

(4) 多个细缆网段可通过中继器连接。但一个完整的 10Base-2 以太网最多能包括 4 个中继器,从而最多连接 5 个网络段,使整个网络的范围达到 925m。

从以上的技术规范可见,10Base-2 的规模和速度受到很大的限制,因此这种网络只适合小范围使用,并不能很好地适用于大型局域网。使用 10Base-2 的主要优点是它的低成本和易安装性。

3. 10Base-T(双绞线网络)

T 表示双绞线星形网,使用 RJ-45 连接器连接。10Base-T 的出现是局域网发展史上的一个非常重要的里程碑,它为以太网在局域网中的统治地位奠定了基础。

10Base-T 标准使用了集线器(Hub)代替公用电缆,这种协议使得设置和维护网络容易些。10Base-T 使用双绞线作为传输介质,在连接距离远的集线器和建筑物时,它可以是一种选择。图 5.12 显示了 10Base-T 的示意图。

10Base-T 要求的技术规范如下。

(1) 每个节点使用 RJ-45 连接器,在工作站端用于连接网络电缆和网络接口卡,在网络端用于连接电缆和集线器。

(2) 一个 10Base-T 段跨越的最大距离是 100m。

(3) 可以通过集线器或交换机扩展网络的覆盖范围,但同样最多允许连接 5 个连续的网络段。

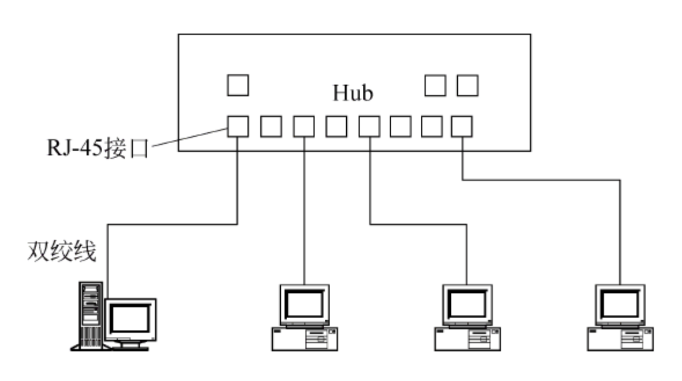


图 5.12 10Base-T 以太网

10Base-T 一般可以适用于大多数工作场合，对于工作节点的物理分布没有什么要求，另外由于其出色的可靠性和易维护性，使得这种网络技术成为现在局域网设计的首选方案。

由于物理结构的问题，在使用非屏蔽双绞线电缆时，常见的一种问题是干扰问题。双绞线按逆时针方向扭在一起的原因，就是为了消除两条信号线之间的串扰。另外，10Base-T 技术通过在电线上使用噪声平衡和滤波技术来抵消它附近的信号源，如电动机、动力线以及雷达产生的电磁干扰。

10Base-T 以太网上的节点连接到星形结构的中心集线器或中继器上。作为一种典型的星形拓扑结构，单根网络电缆仅仅连接两个设备，这使得 10Base-T 网络比起使用总线型拓扑结构的 10Base-2 或 10Base-5 具有更好的容错性。由于每个设备都独立地连接到局域网上，因此 10Base-T 能够更轻易地隔离故障，从而更容易进行故障检修。

表 5.1 给出了常见以太网物理层标准之间的比较。尽管不同的以太网在物理层存在较大的差异，但它们在数据链路层都是采用 CSMA/CD 作为介质访问控制协议的，并且在 MAC 子层使用统一的 IEEE 802.3 帧格式。所以 10Base-T 网络与 10Base-2、10Base-5 是相互兼容的。事实上，即使在以太网后来的发展中，以太网技术也仍然保留了这种标准的帧格式，从而使得所有的以太网系列技术之间能够相互兼容。

表 5.1 IEEE 802.3 以太网的基本特性

特　　性	10Base-5	10Base-2	10Base-T	10Base-F
速率/Mbps	10	10	10	10
传输方法	基带	基带	基带	基带
最大网段长度/m	500	185	100	2000
站间最小距离/m	2.5	0.5		
传输介质	50Ω 粗同轴电缆	50Ω 细同轴电缆	UTP	多模光缆
网络拓扑	总线型	总线型	星形	点对点

5.4.3 快速以太网

速率达到或超过 100Mbps 的以太网称为快速以太网。

快速以太网技术 100Base-T 是由 10Base-T 标准以太网发展而来的，主要解决网络带宽在局域网应用中的瓶颈问题。其协议标准为 1995 年颁布的 IEEE 802.3u，可支持 100Mbps

的数据传输速率,并且与10Base-T一样可支持共享式与交换式两种使用环境,在交换式以太网环境中可以实现全双工通信。IEEE 802.3u在MAC子层仍采用CSMA/CD作为介质访问控制协议,并保留了IEEE 802.3的帧格式。但是,为了实现100Mbps的传输速率,在物理层作了一些重要的改进。例如,在编码上,采用了效率更高的编码方式。

传统以太网采用曼彻斯特编码,其优点是具有自带时钟特性,能够将数据和时钟编码在一起,但其编码效率只能达到1/2,即在具有20Mbps传输能力的介质中,只能传输10Mbps的信号。快速以太网采用4B/5B编码。

1. 快速以太网的体系结构

图5.13给出了IEEE 802.3u协议的体系结构,对应于OSI参考模型的数据链路层协议和物理层协议。

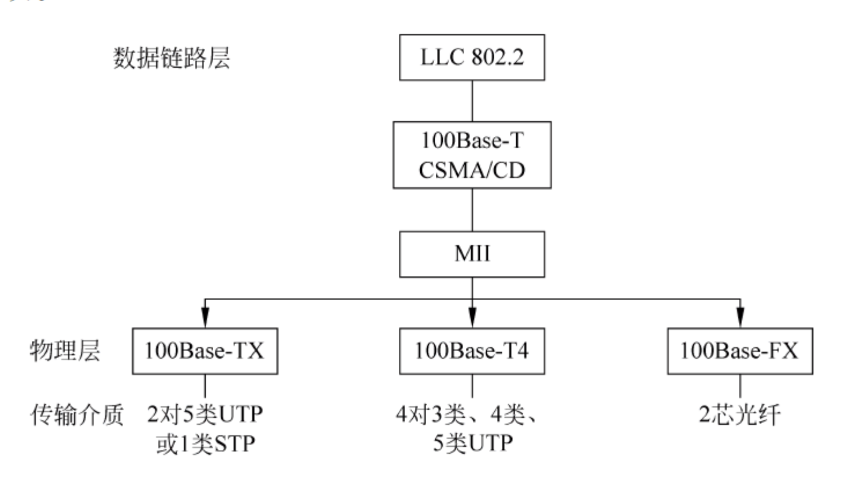


图5.13 100Base-T协议结构

2. 100Base-T物理层

从图5.13中可以看出,100Base-T定义了3种不同的物理层协议。表5.2给出了这3种物理层标准的对比。为了屏蔽下层不同的物理细节,为MAC和高层协议提供了一个100Mbps传输速率的公共透明接口,快速以太网在物理层和MAC子层之间还定义了一种独立于介质种类的介质无关接口(Medium Independent Interface,MII),该接口可以支持上面3种不同的物理层介质标准。

表5.2 100Base-T的3种不同的物理层协议

物理层协议	线缆类型	线缆对数	最大分段长度	编码方式	优点
100Base-T4	3/4/5类UTP	4对	100m	8B/6T	3类UTP
100Base-TX	5类UTP/RJ-45接头 1类STP/DB-9接头	2对	100m	4B/5B	全双工
100Base-FX	62.5μm单模/125μm多模光纤	2对	2000m	4B/5B	全双工长距离

1) 100Base-TX

100Base-TX介质规范基于ANSI TP-PMD物理介质标准。100Base-TX介质接口在两对双绞线电缆上运行,其中一对用于发送数据,另一对用于接收数据,由于ANSI TP-PMD规范既包括屏蔽双绞线电缆,也包括非屏蔽双绞线电缆,所以100Base-TX介质接口支持两对5类以上非屏蔽双绞线电缆和两对1类屏蔽双绞线电缆。

100Base-TX 链路与介质相关的接口有两种：对非屏蔽双绞线电缆，MDI 连接器必须是兼容 5 类及 5 类以上的 8 脚 RJ-45 连接器；对屏蔽双绞线电缆，MDI 连接器必须是 IBM 的 STP 连接器，使用屏蔽 DB-9 型连接器。

(1) 5 类 UTP 及 5 类以上 UTP：100Base-TX UTP 介质接口使用两对 MDI 连接器线来将信号传出和传入网络介质，这意味着 RJ-45 连接器 8 个管脚中的 4 个是被占用的。为使串音和可能的信号失真最小，另外 4 条线不应传输任何信号。每对的发送和接收信号是极化的，一条线传输正(+)信号，而另一条线传输负(－)信号。对 RJ-45 连接器，正确的配线对分配是管脚[1，2]和管脚[3，6]。应尽量在 MDI 管脚分配中使用正确的彩色编码线对。如表 5.3 所示即为 100Base-TX 的 UTP MDI 连接器管脚分配表。

表 5.3　100Base-TX 的 UTP MDI 连接器管脚分配表

管脚号	信号名	电缆颜色
1	发送_	白色/橙色
2	发送_	橙色/白色
3	接收_	白色/绿色
4	保留	
5	保留	
6	接收_	绿色/白色
7	保留	
8	保留	

(2) 1 类 STP：100Base-TX 标准也支持特征阻抗为 150Ω 的屏蔽双绞线电缆。屏蔽双绞线电缆使用 D 型连接器并按 ANSI TP-PMD 对屏蔽双绞线架设的规范来布线。在 DB-9 连接器上正确的配线对分配是管脚[1，6]和管脚[5，9]。如表 5.4 所示即为 100Base-TX 的 STP MDI 连接器管脚分配表。

表 5.4　100Base-TX 的 STP MDI 连接器管脚分配表

管脚号	信号名	电缆颜色
1	接收_	橙色
2	保留	
3	保留	
4	保留	
5	发送_	红色
6	接收_	黑色
7	保留	
8	保留	
9	发送_	绿色
10	底盘	电缆外壳

(3) 100Base-T 交叉布线：当两个节点在网段上连到一起时，一个 MDI 连接器的发送对连到第二个节点的 MDI 的接收对。当两个节点连到一起用于单机应用时，必须提供一条

外部交叉电缆，将电缆一端8脚RJ-45连接器上的发送管脚连到电缆另一端8脚RJ-45连接器上的接收管脚。在多个节点连到一个集线器或交换机端口的实现中，交叉布线是在集线器或交换机端口内部完成的，这使得直连电缆能用于各个节点和集线器或交换机端口之间。如表5.5所示即为100Base-TX交叉连接管脚分配表。

表5.5 100Base-TX交叉连接管脚分配表

管脚号	5类UTP电缆		1类STP电缆	
	无交叉信号名	交叉信号名	无交叉信号名	交叉信号名
1	发送_	接收_	接收_	发送_
2	发送_	接收_	保留	保留
3	接收_	发送_	保留	保留
4	保留	保留	保留	保留
5	保留	保留	发送_	接收_
6	接收_	发送_	接收_	发送_
7	保留	保留	保留	保留
8	保留	保留	保留	保留
9	N/A	N/A	发送_	接收_
10	N/A	N/A	底盘	底盘

(4) 100Base-TX电缆配置：快速以太网的电缆配置安装应符合EIA/TIA-568标准，它描述了接线箱和网络节点之间准确的电缆长度。这一段长度的电缆称为网段，并在以太网规范中被定义为链段。链段正式定义为连接两个且仅仅连接两个MDI的点到点的介质。100Base-TX规范允许两个DTE或DTE与交换端口的链路之间的链段最大长度为100m。

2) 100Base-FX

光缆是100Base-FX指定支持的一种介质，而且容易安装，重量轻，体积小，灵活性好，不受EMI干扰。

100Base-FX标准指定了两条多状态光纤，一条用于发送数据，一条用于接收数据。当工作站的NIC以全双工模式运行时能超过2km。光缆可分为两类：多模光缆和单模光缆。

(1) 多模光缆：这种光缆为62.5/125μm，采用基于LED的收发器将波长为820nm的光信号发送到光纤上。当连在两个设置为全双工模式的交换机端口之间时，支持的最大距离为2km。

(2) 单模光缆：这种光缆为9/125μm，采用基于激光的收发器将波长为1300nm的光信号发送到光纤上。单模光缆率损耗小，较之多模光缆能使光信号传输到更远的距离。

3) 100Base-T4

100Base-T4是100Base-T标准中唯一全新的PHY标准。100Base-T4标准是用来帮助那些已经安装了第3类或第4类电缆的用户的。

100Base-T4链路与介质相关的接口是基于3类、4类、5类非屏蔽双绞线电缆。100Base-T4标准使用4对线。用于100Base-T的RJ-45连接器也可用于100Base-T4。4对中的3对用于一起发送数据，同时第4对用于冲突检测。每对线都是极化的，每对中的一条线传输正(+)信号，而另一条线传输负(—)信号。如表5.6所示即为100Base-T4 UTP MDI管脚分配表。

表 5.6　100Base-T4 UTP MDI 管脚分配表

管脚号	信号名	电缆编码
1	TX_D1_	白色/橙色
2	TX_D1_	橙色/白色
3	RX_D2_	白色/绿色
4	BI_D3_	蓝色/白色
5	BI_D3_	白色/蓝色
6	RX_D2_	绿色/白色
7	BI_D4_	白色/棕色
8	BI_D4_	棕色/白色

（1）100Base-T4 交叉布线：当两个节点在网段上连接到一起时，一个 MDI 连接器的发送对连接第二个节点 MDI 的接收对。当两个节点连到一起用于单机应用时，必须提供一条外部交叉电缆，将电缆一端 8 脚 RJ-45 连接器上的发送管脚连到电缆另一端 8 脚 RJ-45 连接器上的接收管脚。在多个节点连到一个集线器或交换机端口的实现中，交叉布线是在集线器或交换机端口内部完成的，这使得直连电缆能用于各个节点和集线器或交换机端口之间。如表 5.7 所示即为 100Base-T4 交叉连接管脚分配表。

表 5.7　100Base-T4 交叉连接管脚分配表

管脚号	信号名	管脚号	信号名
1	TX_D1_	1	RX_D2_
2	TX_D1_	2	RX_D2_
3	RX_D2_	3	TX_D1_
4	BI_D3_	4	BI_D4_
5	BI_D3_	5	BI_D4_
6	RX_D2_	6	TX_D1_
7	BI_D4_	7	BI_D3_
8	BI_D4_	8	BI_D3_

（2）8B6T 编码方式：8B6T 编码方式有效地将字节的每位映射到一个称为 6T 代码组的 6b 三进制符号内，这就是 8B6T。6T 代码组散开到 3 个发送组上，有效的数据传输速率为 100Mbps 的 1/3，即 33.3Mbps。每对线上的三进制符号的传输速率是 33.3Mbps 的 6/8，即 25 MHz，与 MII 时钟的频率相同，因此 100Base-T4 PHY 中不需要 PLL（锁相回路）。每对上发送的三进制符号可以有 3 个值，与有两个值的二进制符号不一样。

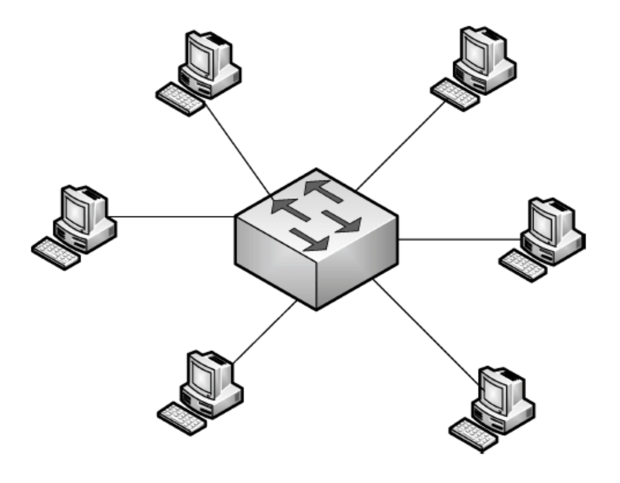


图 5.14　100Base-T 快速以太网组网举例

图 5.14 给出了一个采用 100Mbps 交换机进行组网的快速以太网的例子。由于高速以太网是从 10Base-T 发展而来的，并且保留

了 IEEE 802.3 的帧格式，所以 10Mbps 以太网可以非常平滑地过渡为 100Mbps 快速以太网。

快速以太网的最大优点是结构简单、实用、成本低并易于普及。目前主要用于快速桌面系统，也有少量被用于小型园区网络的主干。

5.4.4 千兆位以太网技术

随着多媒体技术、高性能分布计算和视频应用等的不断发展，用户对局域网的带宽提出了越来越高的要求；同时，100Mbps 快速以太网也要求主干网、服务器一级的设备要有更高的带宽。在这种需求背景下人们开始酝酿速度更高的以太网技术。1996 年 3 月 IEEE 802 委员会成立了 IEEE 802.3z 工作组，专门负责千兆位以太网及其标准，并于 1998 年 6 月正式公布关于千兆位以太网的标准。

千兆位以太网标准是对以太网技术的再次扩展，其数据传输速率为 1000Mbps 即 1Gbps，因此也称吉比特以太网。千兆位以太网基本保留了原有以太网的帧结构，所以向下和以太网与快速以太网完全兼容，从而原有的 10Mbps 以太网或快速以太网可以方便地升级到千兆位以太网。千兆位以太网标准实际上包括支持光纤传输的 IEEE 802.3z 和支持铜缆传输的 IEEE 802.3ab 两大部分。图 5.15 给出了千兆位以太网的协议结构。

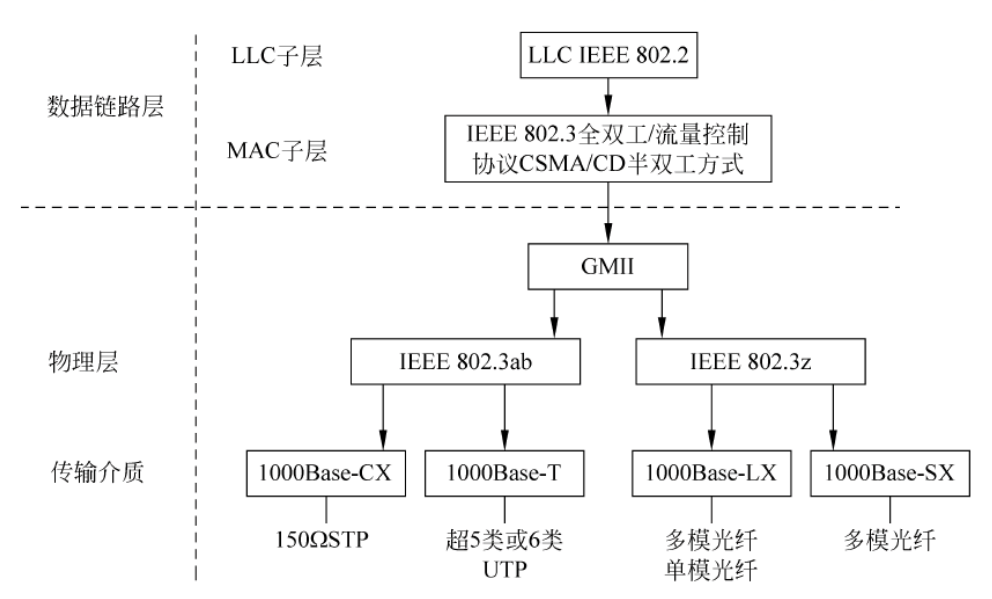


图 5.15 标准的千兆位以太网协议体系

从图 5.15 可以看出，千兆位以太网的物理层包括 1000Base-SX、1000Base-LX、1000Base-CX 和 1000Base-T 4 个协议标准。

1. 1000Base-SX 标准

1000Base-SX 采用芯径为 62.5μm 和 50μm 的多模光纤，工作波长为 850nm，传输距离为 260m 和 525m。数据编码方法为 8B/10B，适用于作为大楼网络系统的主干通路。

2. 1000Base-LX 标准

1）多模光纤

1000Base-LX 可采用芯径为 50μm 和 62.5μm 的多模光纤，工作波长为 850nm，传输距

离为 550m,数据编码方法为 8B/10B,适用于作为大楼网络系统的主干通路。

2) 单模光纤

1000Base-LX 可采用芯径为 9μm 的单模光纤,工作波长为 1300nm 或 1550nm,数据编码方法为 8B/10B,适用于校园或城域主干网。

3. 1000Base-CX 标准

1000Base-CX 标准采用 150Ω 平衡屏蔽双绞线(STP),传输距离为 25m,传输速率为 1.25Gbps,数据编码方法为 8B/10B,适用于集群网络设备的互联,例如,机房内连接网络服务器。

4. 1000Base-T 标准

1000Base-T 采用 4 对 5 类 UTP 双绞线,传输距离为 100m,传输速率为 1Gbps,主要用于结构化布线中同一层建筑的通信,从而可以利用以太网或快速以太网已铺设的 UTP 电缆,也可被用作大楼内的网络主干。

在千兆位以太网的 MAC 子层,除了支持以往的 CSMA/CD 协议外,还引入了全双工流量控制协议。其中,CSMA/CD 协议用于共享信道的争用问题,即支持以集线器作为星形拓扑中心的共享以太网组网;全双工流量控制协议适用于交换机到交换机或交换机到站点之间的点—点连接,两点间可以同时进行发送与接收,即支持以交换机作为星形拓扑中心的交换以太网组网。

与快速以太网相比,千兆位以太网有其明显的优点。千兆位以太网的速度 10 倍于快速以太网,但其价格只有快速以太网的 2~3 倍,即千兆位以太网具有更高的性能价格比。而且从现有的传统以太网与快速以太网可以平滑地过渡到千兆位以太网,并不需要掌握新的配置、管理与排除故障技术。千兆位以太网的主要优点如下。

(1) 简易性:千兆位以太网保持了经典以太网的技术原理、安装实施和管理维护的简易性,这是千兆位以太网成功的基础之一。

(2) 技术过渡的平滑性:千兆位以太网保持了经典以太网的主要技术特征,采用 CSMA/CD 介质管理协议,采用相同的帧格式及帧的大小,支持全双工、半双工工作方式,以确保平滑过渡。

(3) 网络可靠性:保持经典以太网的安装、维护方法,采用中央集线器和交换机的星形结构和结构化布线方法,以确保千兆位以太网的可靠性。

(4) 可管理性和可维护性:采用简单网络管理协议(SNMP)即经典以太网的故障查找和排除工具,以确保千兆位以太网的可管理性和可维护性。

(5) 网络成本:包括设备成本、通信成本、管理成本、维护成本及故障排除成本。由于继承了经典以太网的技术,使千兆位以太网的整体成本下降。

(6) 支持新应用与新数据类型:计算机技术和应用的发展,出现了许多新的应用模式,对网络提出了更高的要求。千兆位以太网具有支持新应用与新数据类型的高速传输能力。

目前,千兆位以太网主要被用于园区或大楼网络的主干中,但也有的被用于有非常高带宽要求的高性能桌面环境中。图 5.16 给出了一个将千兆位以太网用于网络主干,将快速以太网或 10Mbps 以太网用于桌面环境的网络示意图。该网络采用了典型的层次化网络设计方法。

图 5.16 中，最下面一层由 10Mbps 以太网交换机加上 100Mbps 上行链路组成；第 2 层由 100Mbps 以太网交换机加 1000Mbps 上行链路组成；最高层由千兆位以太网交换机组成。通常将面向用户连接或访问网络的层称为接入层(Access Layer)，而将网络主干层称为核心层(Core Layer)，将连接接入部分和核心部分的层称为分布层或汇聚层(Distribution Layer)。

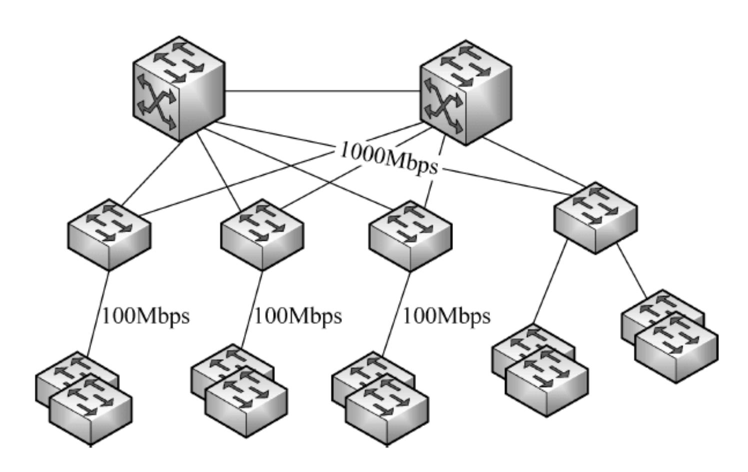


图 5.16　千兆位以太网的应用举例

5.4.5　万兆位以太网技术

在以太网技术中，快速以太网是一个里程碑，确立了以太网技术在桌面的统治地位。随后出现的千兆位以太网更是加快了以太网的发展。然而以太网主要是在局域网中占绝对优势，在很长一段时间中，由于带宽以及传输距离等原因，人们普遍认为以太网不能用于城域网，特别是在汇聚层以及骨干层。1999 年年底成立了 IEEE 802.3ae 工作组进行万兆位以太网技术(10Gbps)的研究，并于 2002 年正式发布 IEEE 802.3ae 10GE 标准。万兆位以太网不仅再度扩展了以太网的带宽和传输距离，更重要的是使得以太网从局域网领域向城域网领域渗透。

正如 1000Base-X 和 1000Base-T(千兆位以太网)都属于以太网一样，从速度和连接距离上来说，万兆位以太网是以太网技术自然发展中的一个阶段。但是，它是一种只适用于全双工模式，并且只能使用光纤的技术。

1. 万兆位以太网的技术特色和显著特征

万兆位以太网相对于千兆位以太网拥有绝对的优势和特点。

(1) 表现在物理层上。万兆位以太网是一种只采用全双工与光纤的技术，其物理层(PHY)和 OSI 参考模型的第 1 层(物理层)一致，它负责建立传输介质(光纤或铜线)和 MAC 子层的连接，MAC 子层相当于 OSI 参考模型的第 2 层(数据链路层)。在网络的结构模型中，把 PHY 进一步划分为物理介质相关子层(PMD)和物理编码子层(PCS)。光学转换器属于 PMD 子层。PCS 子层由信息的编码方式(如 64B/66B)、串行或多路复用等功能组成。

(2) 万兆位以太网技术基本承袭了以太网、快速以太网及千兆位以太网技术，因此在用户普及率、使用方便性、网络互操作性及简易性上皆占有极大的引进优势。在升级到万兆位以太网解决方案时，用户不必担心既有的程序或服务会受到影响，升级的风险非常低，同时

在未来升级到40Gbps甚至100Gbps时都将有很明显的优势。

(3) 万兆位标准意味着以太网将具有更高的带宽(10Gbps)和更远的传输距离(最长传输距离可达40km)。

(4) 在企业网中采用万兆位以太网可以最好地连接企业网骨干路由器,这样大大简化了网络拓扑结构,提高了网络性能。

(5) 万兆位以太网技术提供了更多的更新功能,大大提升了QoS。因此,能更好地满足网络安全、服务质量、链路保护等多个方面需求。

(6) 随着网络应用的深入,WAN/MAN与LAN融合已经成为大势所趋,各自的应用领域也将获得新的突破,而万兆位以太网技术让工业界找到了一条能够同时提高以太网的传输速率、可操作距离和连通性的途径,万兆位以太网技术的应用必将为“三网”发展与融合提供新的动力。

万兆位以太网还有以下十分明显的应用特征。

(1) 万兆位以太网结构简单,管理方便,价格低廉。由于没有采用访问优先控制技术,简化了访问控制的算法,从而简化了网络的管理,并降低了部署的成本,因而得到了广泛的应用。

(2) 过去有时需采用数个千兆位捆绑在一起以满足交换机互联所需的高带宽,因而浪费了更多的光纤资源,现在可以采用万兆位互联,甚至4个万兆位捆绑互联,达到40Gbps的宽带水平。

(3) 采用万兆位以太网,网络管理者可以用实时方式,也可以用历史累积方式轻松地看到第2层至第7层的网络流量。允许“永远在线”监视,能够鉴别干扰或入侵监测,发现网络性能瓶颈,获取计费信息或呼叫数据记录,从网络中获取商业智能。

(4) 以太网的可平滑升级保护了用户的投资,以太网的改进始终保持向前兼容,使得用户能够实现无缝升级,一方面不需要额外的投资升级上层应用系统,也不影响原来的业务部署和应用。

2. 万兆位以太网技术介绍

如图5.17所示为IEEE 802.3ae万兆位以太网技术标准的体系结构。

1) 物理层

在物理层,万兆位以太网的IEEE 802.3ae标准只支持光纤作为传输介质,但提供了两种物理连接(PHY)类型。一种是提供与传统以太网进行连接的速率为10Gbps的LAN物理层设备,即LAN PHY;另一种是提供与SDH/SONET进行连接的速率为9.58464Gbps的WAN物理层设备,即WAN PHY。通过引入WAN PHY,提供了以太网帧与SONETOC-192帧结构的融合,WAN PHY可与OC-192、SDH/SONET设备一起运行,从而在保护现有网络投资的基础上,能够在不同地区通过SONET城域网提供端到端的以太网连接。

每种PHY分别可使用10GBase-S(850nm短波)、10GBase-L(1310nm长波)和10GBase-E(1550nm长波)3种规格,最大传输距离分别为300m、10km、40km。

在物理拓扑上,万兆位以太网既支持星形连接或扩展星形连接,也支持点到点连接及星形连接与点到点连接的组合,在万兆位以太网的MAC子层,已不再采用CSMA/CD机制,只支持全双工方式。事实上,尽管在千兆位以太网协议标准中提到了对CSMA/CD的支持,但基本上只采用全双工方式,而不再采用共享带宽方式。另外,其继承了802.3以太网

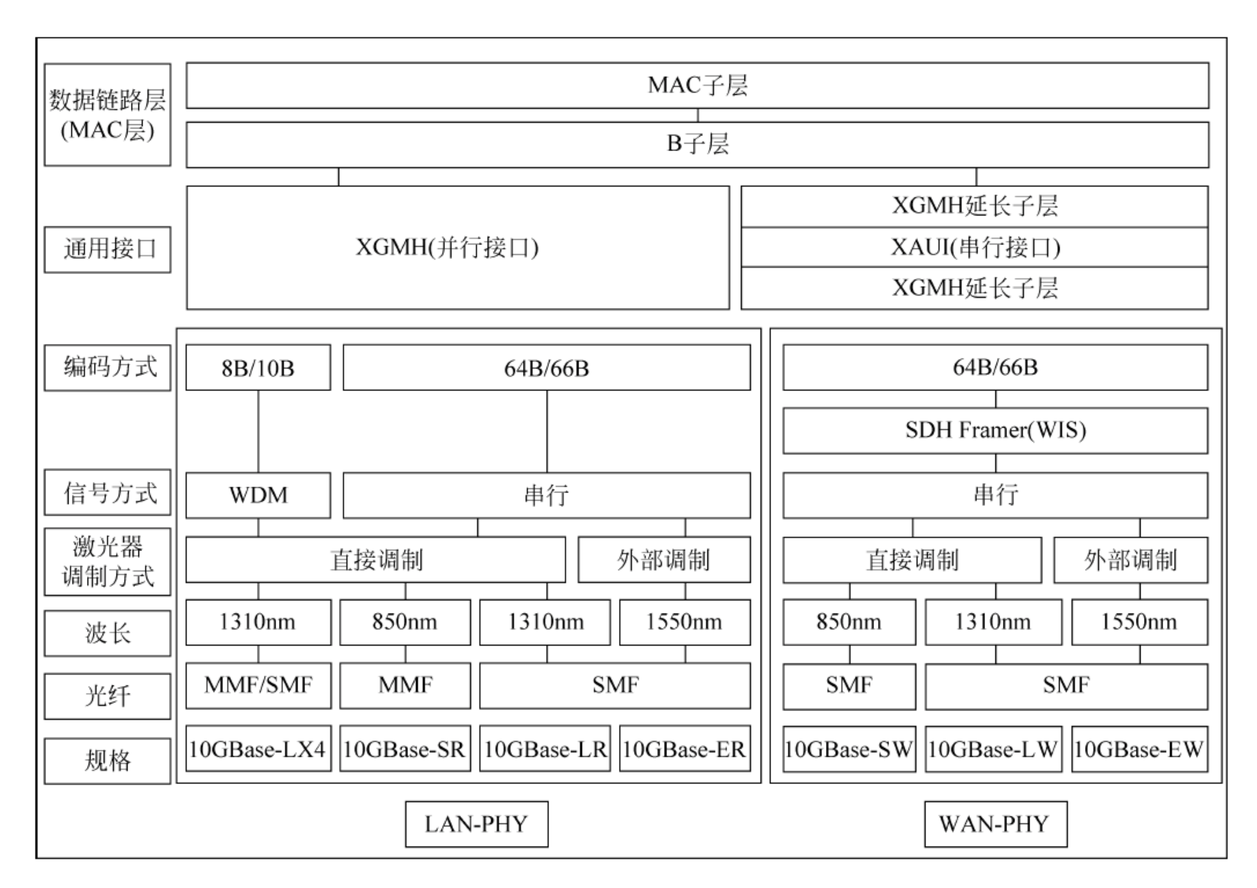


图 5.17 IEEE 802.3ae 体系结构

的帧格式和最大/最小帧长度，从而能充分兼容已有的以太网技术，进而降低了对现有以太网进行万兆位升级的风险。

(1) 10Gbps 串行物理媒介层。万兆位以太网支持 5 种接口，分别是 1550nm LAN 接口、1310nm 宽频波分多路复用(WWDM)LAN 接口、850nm LAN 接口、1550nm WAN 接口和 1310nm WAN 接口。每种接口都有其对应的最便宜的传输介质，传输距离也不同，如表 5.8 所示。

表 5.8 10Gbps 串行物理媒介层

名　称	描　述	传输介质	传输距离
10GBase-SR	850nm LAN 接口	50/125μm 多模光纤	65m
10GBase-LR	1310nm LAN 接口	62.5/125μm 多模光纤	300m
10GBase-ER	1550nm LAN 接口	50/125μm 多模光纤	240m
10GBase-LW	1310nm LAN 接口	单模光纤	10km
10GBase-EW	1550nm LAN 接口	单模光纤	40km

(2) PMD(物理介质相关)子层。PMD 子层的功能是支持在 PMA 子层和介质之间交换串行化的符号代码位。PMD 子层将这些电信号转换成适合于在某种特定介质上传输的形式。PMD 是物理层的最低子层，标准中规定物理层负责从介质上发送信号和接收信号。

(3) PMA(物理介质接入)子层。PMA 子层提供了 PCS 子层和 PMD 子层之间的串行

化服务接口。和 PCS 子层的连接称为 PMA 服务接口。另外,PMA 子层还从接收位流中分离出用于对接收到的数据进行正确的符号对齐(定界)的符号定时时钟。

(4) WIS(广域网接口)子层。WIS 子层是可选的物理子层,可用在 PMA 与 PCS 之间,产生适配 ANSI 定义的 SONET STS-192c 传输格式或 ITU 定义 SDH VC-4-64c 容器速率的以太网数据流。该速率数据流可以直接映射到传输层而不需要高层处理。

(5) PCS(物理编码)子层。PCS 子层位于协调子层(通过 GMII)和物理介质接入(PMA)子层之间。PCS 子层完成将经过完善定义的以太网 MAC 功能映射到现存的编码和物理层信号系统的功能上去。PCS 子层和上层 RS/MAC 的接口由 XGMII 提供,与下层 PMA 接口使用 PMA 服务接口。

(6) RS(协调子层)和 XGMII(10Gbps 介质无关接口)。协调子层的功能是将 XGMII 的通路数据和相关控制信号映射到原始 PLS 服务接口定义(MAC/PLS)接口上。XGMII 接口提供了 10Gbps 的 MAC 和物理层间的逻辑接口。XGMII 和协调子层使 MAC 可以连接到不同类型的物理介质上。

2) 传输介质层

IEEE 802.3ae 目前支持 9/125μm 单模、50/125μm 多模和 62.5/125μm 多模 3 种光纤,而对电接口的支持规范 10GBase-CX4 目前正在讨论之中,尚未形成标准。

3) 数据链路层

IEEE 802.3ae 继承了 802.3 以太网的帧格式和最大/最小帧长度,支持多层星形连接、点到点连接及其组合,充分兼容已有应用,不影响上层应用,进而降低了升级风险。与传统的以太网不同,IEEE 802.3ae 仅仅支持全双工方式,不支持单工和半双工方式,不采用 CSMA/CD 机制;IEEE 802.3ae 不支持自协商,可简化故障定位,并提供广域网物理层接口。

3. 以太网的应用和展望

1) 万兆位以太网的应用场合

随着千兆到桌面的日益普及,万兆位以太网技术将会在汇聚层和骨干层上广泛应用。从目前网络现状看,万兆位以太网最先应用的场合包括教育网、数据中心出口和城域网骨干。

(1) 在教育网的应用。随着高校多媒体网络教学、数字图书馆等应用的展开,高校校园网将成为万兆位以太网的重要应用场合。利用 10GE 高速链路构建校园网的骨干链路和各分校区与本部之间的连接,可实现端到端的以太网访问,进而提高传输速率,有效地保证远程多媒体教学和数字图书馆等业务的开展。

(2) 在数据中心出口的应用。随着服务器纷纷采用千兆位链路连接网络,汇聚这些服务器的上行带宽将逐渐成为业务瓶颈,使用 10GE 高速链路可为数据中心出口提供充分的带宽保障。

(3) 在城域网的应用。随着城域网建设的不断深入,各种内容业务(如流媒体视频应用、多媒体互动游戏)纷纷出现,这些对城域网的带宽提出更高的要求,而传统的 SDH、DWDM 技术作为骨干存在着网络结构复杂、难以维护和建设成本高等问题。在城域网骨干层部署 10GE 可大大地简化网络结构、降低成本、便于维护,通过端到端以太网打造低成本、高性能和具有丰富业务支持能力的城域网。

2）万兆位以太网的特点

万兆位以太网技术提供更加丰富的带宽和处理能力，能够有效地节约用户在链路上的投资，并保持以太网一贯的兼容性、简单易用和升级容易的特点。但是，由于万兆位以太网尚处于发展初期，还存在着一些问题和不足：①在价格方面，目前一个10GE端口的价格是GE端口的100倍左右，尤其是在带宽得不到充分利用的情况下，会造成投资的极大浪费；②万兆位以太网继承了以太网一贯的弱QoS特点，如何进行有保障的区分业务承载的问题仍然没有解决，RPR、MPLS等特性的支持尚不成熟；③10GE要求设备具有强大的处理能力，而目前业界有些厂商推出的10GE端口并未实现真正的线速处理，带宽优势大打折扣。

5.4.6 异步传输模式网络（ATM）

随着人们对语音、图像和数据为一体的多媒体通信需求的日益增加，特别是为了适应今后信息高速公路建设的需要，人们又提出了宽带综合业务数字网（B-ISDN），这是一种全新的通信网络，而B-ISDN的实现需要一种全新的传输模式，即异步传输模式（ATM）。在1990年，国际电报电话咨询委员会（CCITT）正式建议将ATM作为实现B-ISDN的一项技术基础。

为了更好地了解ATM，有必要先对时分多路复用（TDM）和同步传输（STM）作一简单的回顾。TDM即是在一条通信线路上按一定的周期（如125ns）将时间分成称为帧的时间块，而在每一帧中又分成若干时隙，每个时隙可携带相应的用户信息。当某一用户通过呼叫建立起通信后，在此期间，其信号将固定地占用各帧中的某一时隙，直至通信结束，如图5.18所示。

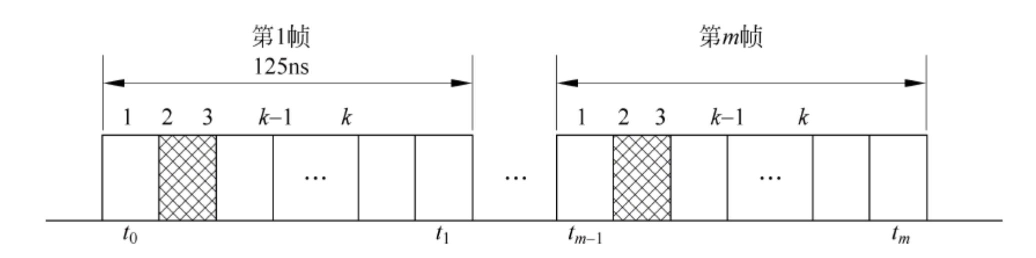


图5.18 同步传输模式示意图

由图5.18可见，对于同步传输，其交换是在固定时隙之间进行的。例如，图5.18中，在输入帧占用第2时隙的某一信号，若在输出帧中占用的也是第2时隙，则这种对应关系是固定不变的，直至相应的通信过程结束。

在这种固定时隙的传输及交换模式中，若在通信过程中的某一时刻用户无数据传递，其固定占用的时隙仍属其所有，尽管此刻处于空闲状态；相反，若其有大量突发性数据要求传送，尽管这有可能造成信号的延时甚至是信元的丢失，也仍只能借助于固定的时隙来传输和交换，这样就造成了很大的浪费。

相比之下，在异步传输模式（ATM）中，其信元传输所占用的时隙并不固定，这也是所谓的统计时分多路复用。另外，在一帧中占用的时隙数也不固定，可以有一个至多个时隙，完全根据当时用户通信的情况而定，而且各时隙之间并不要求连续，纯粹是“见缝插针”。在交

换时,也是类似情况。这个过程如图 5.19 所示。

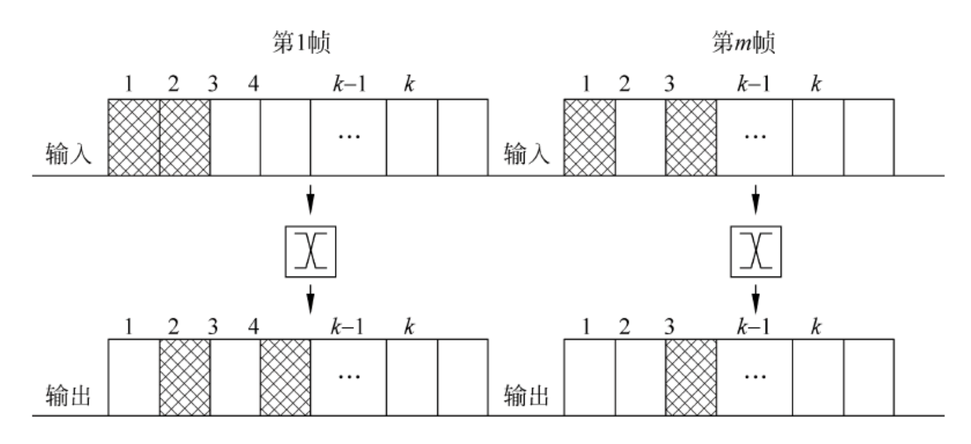


图 5.19　异步传输模式示意图

由于在 ATM 中具有动态分配带宽的特点,可以充分地利用带宽资源,并且能很好地满足传输突发性数据的要求,而不至于出现在 ATM 中延时或信元丢失的情况。

ATM 扩大了网络能力以支持各种各样的应用和帮助网络管理者。使用 ATM 技术作为基础,确保一个与广域网(WAN)的可接受界面,满足 LAN 中高宽带业务需求。传输管理设计保证用户得到他们所需要的服务质量。

5.5　无线局域网

随着无线局域网(Wireless Lan,WLAN)技术的发展,人们越来越深刻地认识到,无线局域网不仅能够满足移动和特殊应用领域网络的要求,还能覆盖有线网络难以涉及的范围。无线局域网以微波、激光与红外线等无线光波作为传输介质,以此来部分或全部代替传统局域网中的同轴电缆、双绞线与光纤,实现了移动计算网络中移动节点的物理层与数据链路层功能,并为移动计算网络提供物理网接口,可以作为传统局域网的有效补充。无线局域网的研究与应用已经成为网络技术一个热点问题。

与有线网络相比,无线局域网具有以下优点。

(1) 安装便捷。一般在网络建设中,施工周期最长、对周边环境影响最大的,就是网络综合布线工程中,往往需要破墙掘地、穿线架管。而无线局域网最大的优势就是免去或减少了网络布线的工作量,一般只要安装一个或多个接入点 AP(Access Point)设备,就可建立覆盖整个建筑或地区的局域网。

(2) 使用灵活。在有线网络中,网络设备的安放位置受网络信息点位置的限制。而一旦无线局域网建成后,在无线网的信号覆盖区域内任何一个位置都可以接入网络。

(3) 节约成本。由于有线网络缺少灵活性,这就要求网络规划者尽可能地考虑未来发展的需要,这就往往导致预设大量利用率较低的信息点。而一旦网络的发展超出了设计规划,又要花费较多费用进行网络改造,而无线局域网可以避免或减少以上情况的发生。

(4) 易于扩展。无线局域网有多种配置方式,能够根据需要灵活选择。这样,无线局域网就能胜任从只有几个用户的小型局域网到上千用户的大型网络,并且能够提供像“漫游

(Roaming)”等有线网络无法提供的特性。

5.5.1 无线局域网的技术标准

目前,支持无线局域网的技术标准主要有蓝牙(Bluetooth)技术、HomeRF 技术以及 IEEE 802.11 系列。

1. 蓝牙技术

蓝牙技术是在 1994 年爱立信为寻找蜂窝电话和 PDA 那样的辅助设备进行通信的廉价无线接口时创立的。它是一种支持设备短距离通信(一般是 10m 之内)的无线电技术。能在包括移动电话、PDA、无线耳机、笔记本电脑、相关外设等众多设备之间进行无线信息交换。蓝牙的标准是 IEEE 802.15,工作在 2.4GHz 频带,带宽为 1Mbps,采用高速跳频(Frequency Hopping,FH)技术和时分多址(Time Division Multiple Access,TDMA)技术。可在近距离内以非常廉价的成本将一些数字化设备(各种移动设备、固定通信设备、计算机及其终端设备、各种数字数据系统,如数字照相机、数字摄像机等,甚至各种家用电器、自动化设备)呈网状链接起来。其有效距离一般为 10m 左右,传输速率约为 1Mbps。

2. HomeRF 技术

HomeRF 无线标准是由 HomeRF 工作组开发的,主要用于家庭无线网络,使计算机与其他电子设备之间实现无线通信的开放性工业标准。HomeRF 使用开放的 2.4GHz 频段,采用跳频扩频(FHSS)技术。其数据通信采用简化的 IEEE 802.11 协议标准,沿用类似于以太网技术中的冲突检测的载波监听多路访问/冲突检测(CSMA/CD)和冲突避免的载波监听多址访问(CSMA/CA)。语音通信采用 DECT(Digital Enhanced Cordless Telephony)标准,使用 TDMA 技术。不过由于 HomeRF 技术没有公开,目前支持的企业不多,而且在抗干扰等方面相对于其他无线技术而言尚有欠缺。

3. 跳频扩展频谱(FHSS)

跳频技术是另外一种扩频技术。跳频的载频受一个伪随机码的控制,在其工作带宽范围内,其频率按随机规律不断改变频率。接收端的频率也按随机规律变化,并保持与发射端的变化规律一致。跳频的高低直接反映跳频系统的性能,跳频越高,抗干扰的性能越好,军用跳频系统可以达到上万跳每秒。实际上移动通信 GSM 系统也是跳频系统。出于成本的考虑,商用跳频系统跳速都较慢,一般在 50 跳/s 以下。由于慢跳跳频系统实现简单,因此低速无线局域网常常采用这种技术。FHSS 局域网支持 1Mbps 数据速率,共有 22 组跳频图案,包括 79 条信道,输出的同步载波经解调后,可获得发送端送来的信息。

与红外线方式比较,使用无线电波作为媒体的 DSSS 和 FHSS 方式,具有覆盖范围大、抗干扰、抗噪声、抗衰减和保密性好的优点。

IEEE 802.11 标准在 MAC 子层采用带冲突避免的载波监听多路访问(Carrier Sense Multiple Access/Collision Avoidance,CSMA/CA)协议。该协议与在 IEEE 802.3 标准中所讨论的 CSMA/CD 协议类似,为了减小无线设备之间在同一时刻同时发送数据导致冲突的风险,IEEE 802.11 引入了称为请求发送/清除发送(RTS/CTS)的机制。即如果发送目的地是无线节点,数据到达基站,该基站将会向无线节点发送一个 RTS 帧,请求一段用来发送数据的专用时间。接收到 RTS 请求帧的无线节点将回应一个 CTS 帧,表示它将中断其

他所有的通信直到该基站传输数据结束。其他设备可监听到传输事件的发生,同时将在此时间段的传输任务向后推迟。这样,节点间传输数据时发生冲突的概率就会大大减少。

5.5.2 无线局域网设备

要组建无线局域网,必须有相应的无线网设备,这些设备主要包括无线网卡、无线访问接入点、无线网桥和天线,几乎所有的无线网产品中都自含无线发射/接收功能。

无线网卡在无线局域网中的作用相当于有线网卡在有线局域网中的作用。按无线网卡的总线类型可分为适用于台式机的 PCI 接口的无线网卡,适用笔记本电脑的 PCMCIA 接口的无线网卡,笔记本电脑和台式机均适用的 USB 接口的无线网卡。

无线访问接入点(AP)则是在无线局域网环境中,进行数据发送和接收的集中设备,相当于有线网络中的集线器。通常,一个 AP 能够在几十米甚至上百米的范围内连接多个无线用户。AP 可以通过标准的 Ethernet 电缆与传统的有线网络相联,从而可作为无线网络和有线网络的连接点。由于无线电波在传播过程中会不断衰减,导致 AP 的通信范围被限定在一定的范围之内,这个范围被称为微单元。但若采用多个 AP,并使它们的微单元相互有一定范围的重合时,则用户可以在整个无线局域网覆盖区内移动,无线网卡能够自动发现附近信号强度最大的 AP,并通过这个 AP 收发数据,保持不间断的网络连接,这种方式称为无线漫游。

无线网桥主要用于无线局域网或有线局域网之间的互联。当两个局域网无法实现有线连接或使用有线连接存在困难时,就可使用无线网桥实现点对点的连接,在这里无线网桥起到了协议转换的作用。

无线路由器则集成了无线 AP 的接入功能和路由器的第三层路径选择功能。

天线(Antenna)功能则是将信号源发送的信号借由天线本身的特性传送至远处。天线一般有所谓定向性(Uni-directional)与全向性(Omni-directional)之分,前者较适合于长距离使用,而后者则较适合于区域性之应用。例如,若要将在第一栋楼内无线网的范围扩展到一千米甚至数千米以外的第二栋楼,其中的一个方法是在每栋楼上安装一个定向天线,天线的方向相互对准,第一栋楼的天线经过网桥连到有线网上,第二栋楼的天线是接在第二栋楼的网桥上,如此无线网就可接通相距较远的两个或多个建筑物。

5.5.3 无线局域网的组网模式

将以上几种无线局域网设备结合在一起使用,就可以组建出多层次、无线与有线并存的计算机网络。一般来说,无线局域网有两种组网模式,一种是无固定基站的;另一种是有固定基站的。这两种模式各有特点,无固定基站组成的网络称为自组网络,主要用于在便携式计算机之间组成平等状态的网络;有固定基站的网络类似于移动通信的机制,网络用户的便携式计算机通过基站(又称为访问点 AP)连入网络。这种网络是应用比较广泛的网络,一般用于有线局域网覆盖范围的延伸或作为宽带无线互联网的接入方式。

1. 自组网络模式

自组网络(Ad-Hoc)又称对等网络,是最简单的无线局域网结构,是一种无中心的拓扑结构,网络连接的计算机具有平等的通信关系,仅适用于较少数的计算机无线互联(通常是在5台主机以内),如图5.20所示。这些计算机要有相同的工作组名和密码(如果适用)。任何时间,只要两个或更多的无线网接口相互都在彼此的范围之内,它们就可以建立一个独立的网络;可以实现点对点与点对多点连接;自组网络不需要固定设施,是临时组成的网络,非常适合于野外作业和军事领域;组建这种网络,只需在每台计算机中插入一块无线网卡,不需要其他任何设备就可以完成通信。

2. 基础结构网络模式

在具有一定数量用户或是需要建立一个稳定的无线网平台时,一般会采用以AP为中心的模式,将有限的"信息点"扩展为"信息区",这种模式也是无线局域网最为普通的构建模式,即基础结构(Infrastructure)模式,采用固定基站的模式。在基础结构网络中,要求有一个无线固定基站充当中心站,所有站点对网络的访问均由其控制,如图5.21所示。

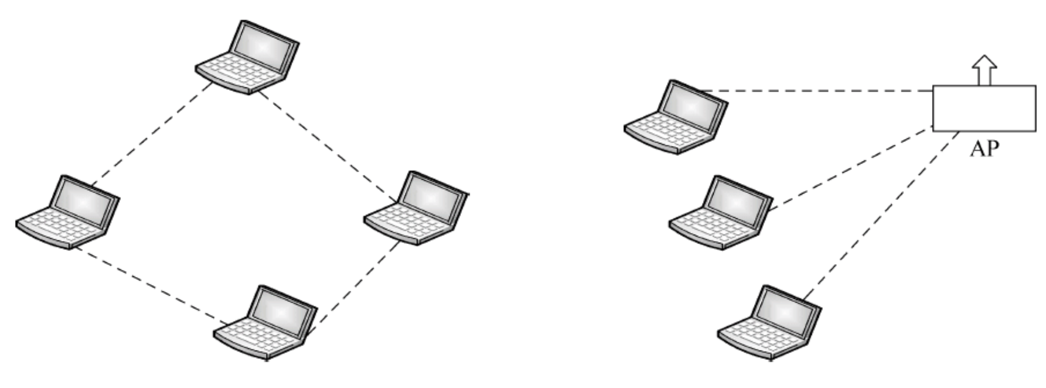


图5.20 对等无线网　　　图5.21 基础结构无线网

在基于AP的无线网中,AP访问点和无线网卡还可针对具体的网络环境调整网络连接速率,如11Mbps的IEEE 802.11b的可使用速率可以调整为1Mbps、2Mbps、5.5Mbps和11Mbps 4种;54Mbps的IEEE 802.11a和IEEE 802.11g的则有54Mbps、48Mbps、36Mbps、24Mbps、18Mbps、12Mbps、11Mbps、9Mbps、6Mbps、5.5Mbps、2Mbps、1Mbps共12个不同速率可动态转换,以发挥相应网络环境下的最佳连接性能。

由于每个站点只需在中心站覆盖范围之内就可与其他站点通信,故网络中站点布局受环境限制较小。

通过无线接入访问点、无线网桥等无线中继设备还可以把无线局域网与有线网连接起来,并允许用户有效地共享网络资源,如图5.22所示。中继站不仅仅提供与有线网的通信,也为网上邻居解决了无线网拥挤的状况。复合中继站能够有效地扩大无线网的覆盖范围,实现漫游功能。有中心网络拓扑结构的弱点是抗毁性差,中心站点的故障容易导致整个网络瘫痪,并且中心站点的引入增加了网络成本。在实际应用中,无线局域网往往与有线主干网结合起来使用。这时,中心站点充当无线局域网与有线主干网的转接器。

3. 无线 Internet 接入

目前,许多公司开始利用WLAN的方式提供移动Internet接入,在宾馆、机场候机大

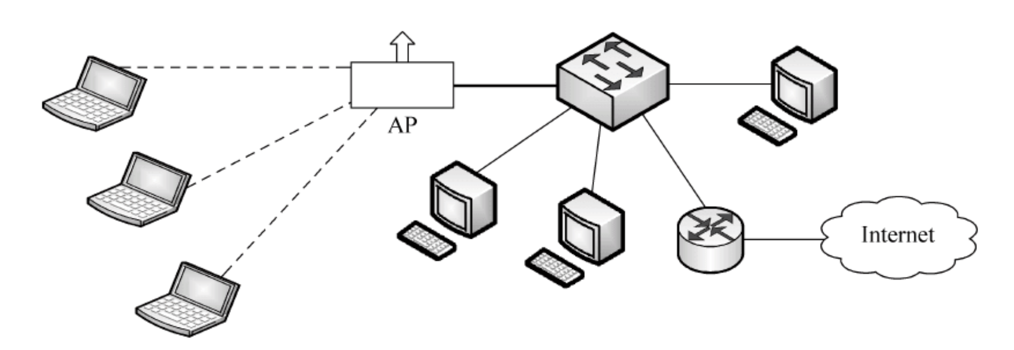


图 5.22　无线与有线的结合实例

厅等地区架设 WLAN，然后通过 DSL 或 FTTX 等方式相结合，为人们提供无线上网的条件。

虽然无线网有诸多优势，但与有线网相比，无线局域网也存在一些不足，如网络速率较慢，价格较高，数据传输的安全性有待进一步提高。因而无线局域网目前主要还是面向那些有特定需求的用户，作为对有线网的一种补充。但也应该看到，随着无线局域网性能价格比的不断提高，它将会在未来发挥更加重要和广泛的作用。

5.6　虚拟局域网

随着以太网技术的普及，以太网的规模也越来越大，从小型的办公环境到大型的园区网络，网络管理变得越来越复杂。第一，在采用共享介质的以太网中，所有节点位于同一冲突域中，同时也位于同一广播域中，即一个节点向网络中某些节点的广播会被网络中所有的节点所接收，造成很大的带宽资源和主机处理能力的浪费。为了解决传统以太网的冲突域问题，采用了交换机来对网段进行逻辑划分。但是，交换机虽然能解决冲突域问题，却不能克服广播域问题。例如，一个 ARP 广播就会被交换机转发到与其相连的所有网段中，当网络上有大量这样的广播存在时，不仅是对带宽的浪费，还会因过量的广播产生广播风暴，当交换网络规模增加时，网络广播风暴问题还会更加严重，并可能因此导致网络瘫痪。第二，在传统的以太网中，同一个物理网段中的节点也就是一个逻辑工作组，不同物理网段中的节点是不能直接相互通信的。这样，当用户由于某种原因在网络中移动但同时还要继续原来的逻辑工作组时，就必然会需要进行新的网络连接乃至重新布线。

为了解决上述问题，虚拟局域网（Virtual Local Area Network，VLAN）应运而生。虚拟局域网是以局域网交换机为基础，通过交换机软件实现根据功能、部门、应用等因素将设备或用户组成虚拟工作组或逻辑网段的技术，其最大的特点是在组成逻辑网时无须考虑用户或设备在网络中的物理位置。VLAN 可以在一台交换机或者跨交换机上实现。

1996 年 3 月，IEEE 802 委员会发布了 IEEE 802.1q VLAN 标准。目前，该标准得到全世界重要网络厂商的支持。

在 IEEE 802.1q 标准中对虚拟局域网（VLAN）是这样定义的：虚拟局域网是由一些局域网网段构成的与物理位置无关的逻辑组，而这些网段具有某些共同的需求。每一个 VLAN 的帧都有一个明确的标识符，指明发送这个帧的工作站是属于哪一个 VLAN。利用

以太网交换机可以很方便地实现虚拟局域网(VLAN)。虚拟局域网其实只是局域网给用户提供的一种服务,而并不是一种新型局域网。

5.6.1　透明和虚拟

1. 透明

如果一个事物或过程是实际的,但它并没有表现出来,看似好像不存在一样,这种属性就是透明。这与日常生活中透明的概念是一样的,生活中如果玻璃擦得很亮很亮,我们可以不必考虑玻璃的存在,透过玻璃看见玻璃那一侧的任何东西,计算机网络中的透明概念也是这样。从用户的角度看,计算机网络通常提供透明的传输,使网络用户可以不必考虑网络的存在而访问网络中的任何资源,如图 5.23 所示,当节点 A 处的主机 H_1 的一个终端用户要访问节点 D 处的主机 H_2 的磁盘时,该用户可用访问一般磁盘文件的方式访问那个文件,就像他访问本地磁盘文件一样。计算机网络的这种特性就是透明性,即对用户来说,网络是透明的,就像擦得很干净的玻璃一样可以不考虑。当然,并非所有的计算机网络对用户都是透明的,有许多网络就要用户在访问他所要求的资源之前,必须提供该资源的地址信息,通过网络建立与该资源的连接,此后,网络对他才是透明的。

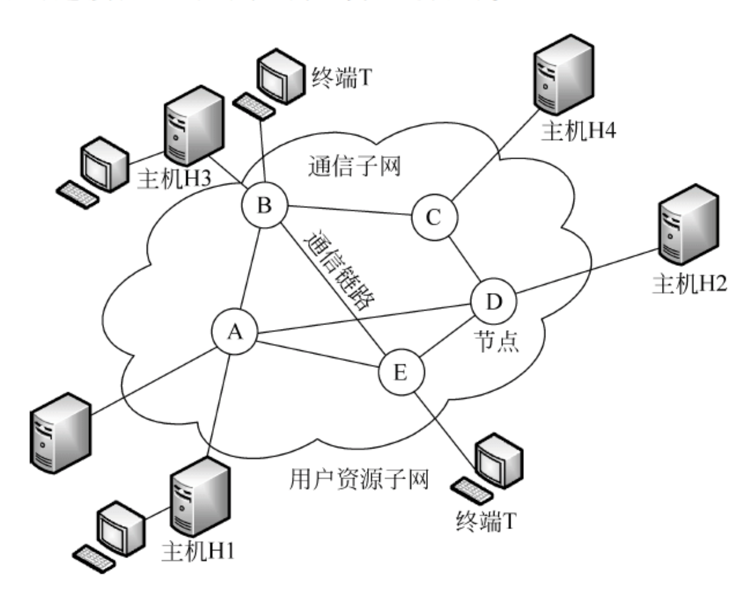


图 5.23　网络的透明传输

2. 虚拟

与透明相对,如果一个事物或过程实际上并不存在,但它却表现出来了,就像实际存在的一样,这种属性就是虚拟。计算机网络技术中常涉及虚拟计算机、虚拟终端、虚拟电路、虚拟存储器、虚拟网关、虚拟功能、虚拟现实等概念。

图 5.24 给出一个关于 VLAN 划分的示例。图中使用了 4 台交换机的网络拓扑结构,有 9 个工作站分配在 3 个楼层中,构成了 3 个局域网,即 LAN1:(A1,B1,C1),LAN2:(A2,B2,C2),LAN3:(A3,B3,C3)。

但这 9 个用户划分为 3 个工作组,也就是说划分为 3 个虚拟局域网(VLAN)。即 VLAN1:(A1,A2,A3),VLAN2:(B1,B2,B3),VLAN3:(C1,C2,C3)。在虚拟局域网上

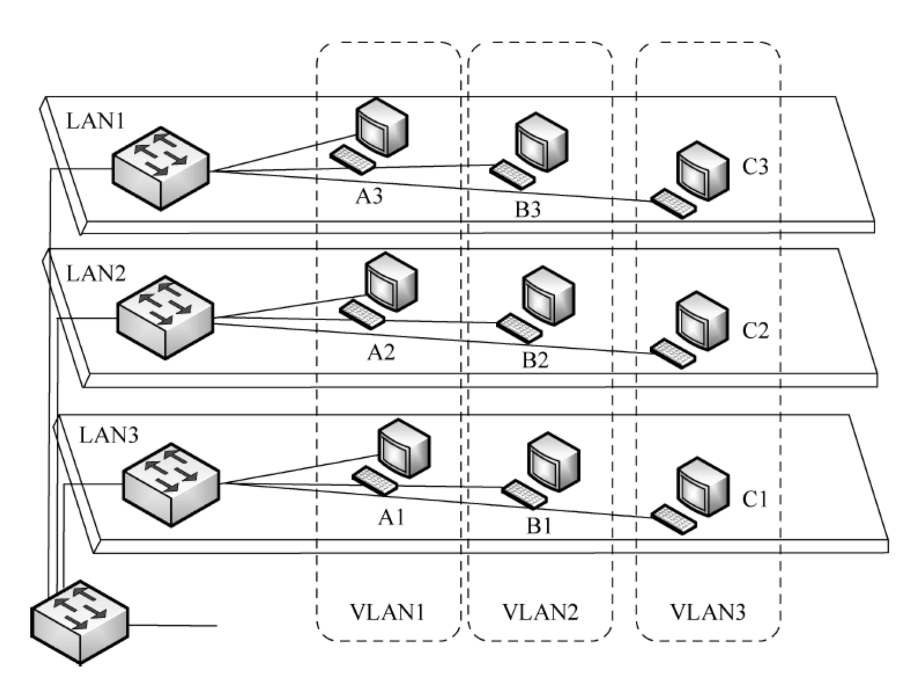


图 5.24　虚拟局域网 VLAN 的示例

的每一个站都可以听到同一虚拟局域网上的其他成员所发出的广播。如工作站 B1、B2、B3 同属于虚拟局域网 VLAN2。当 B1 向工作组内成员发送数据时，B2 和 B3 将会收到广播的信息(尽管它们没有连在同一台交换机上)，但 A1 和 C1 不会收到 B1 发出的广播信息(尽管它们连在同一台交换机上)。

5.6.2　虚拟局域网使用的以太网帧格式

1988 年，IEEE 批准了 802.3ac 标准，这个标准定义了虚拟局域网的以太网帧格式，在传统的以太网的帧格式中插入一个 4B 的标识符，称为 VLAN 标记，用来指明发送该帧的工作站属于哪一个虚拟局域网，如图 5.25 所示。如果还使用传统的以太网帧格式，那么就无法划分虚拟局域网。

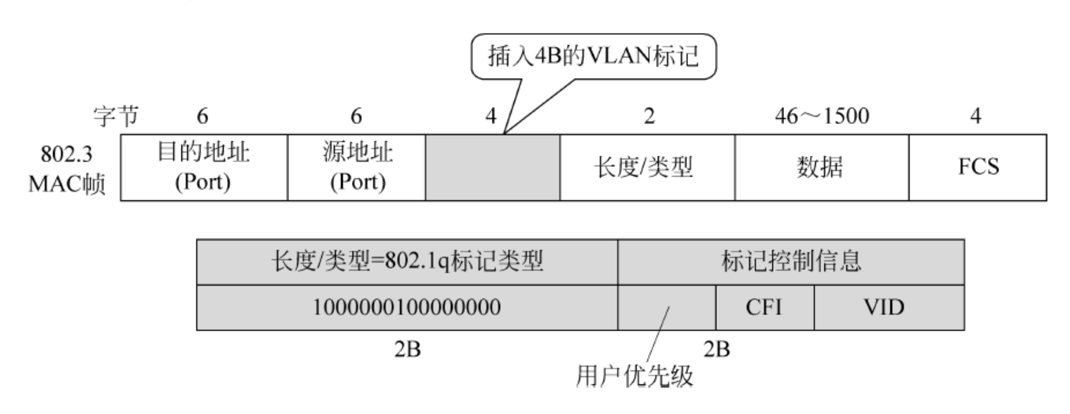


图 5.25　虚拟局域网的以太网帧格式

VLAN 标记字段的长度是 4B，插在以太网 MAC 帧的源地址字段和长度/类型字段之间。VLAN 标记的前两个字节和原来的长度/类型字段的作用一样，但它总是设置为

0x8100(这个数值大于 0x0600,因此不代表长度),称为 802.1q 标记类型。当数据链路层检测到在 MAC 帧的源地址字段后面的长度/类型字段的值是 0x8100 时,就知道现在插入了 4B 的 VLAN 标记。于是就检查该标记的后两个字节的内容。在后面的两个字节中,前 3 个比特是用户优先级字段,接着的一个比特是规范格式指示符(Canonical Format Indicator,CFI),最后的 12 比特是该虚拟局域网的标识符 VID,它唯一地标志这个以太网帧是属于哪一个 VLAN 的。在 IEEE 801.1q 标记(4B)后面的两个字节是以太网帧的长度/类型字段。

因为用于 VLAN 的以太网帧的首部增加了 4B,所以以太网帧的最大长度从原来的 1518B 变为 1522B。

5.6.3 虚拟局域网的优点

采用 VLAN 后,在不增加设备投资的前提下,可在许多方面提高网络的性能,并简化网络的管理。具体表现在以下几方面。

(1) 提供了一种控制网络广播的方法:基于交换机组成的网络的优势在于可提供低时延、高吞吐量的传输性能,但其会将广播包发送到所有互联的交换机、所有的交换机端口、干线连接及用户,从而引起网络中广播流量的增加,甚至产生广播风暴。通过将交换机划分到不同的 VLAN 中,一个 VLAN 的广播不会影响到其他 VLAN 的性能。即使是同一台交换机上的两个相邻端口,只要它们不在同一 VLAN 中,则相互之间也不会渗透广播流量。这种配置方式大大地减少了广播流量,提高了用户的可用带宽,弥补了网络易受广播风暴影响的弱点,同时也是一种比传统的采用路由器在共享集线器间进行网络广播阻隔更灵活有效的方法。

(2) 提高了网络的安全性:VLAN 的数目及每个 VLAN 中的用户和主机是由网络管理员决定的。网络管理员通过将可以相互通信的网络节点放在一个 VLAN 内,或将受限制的应用和资源放在一个安全的 VLAN 内,并提供基于应用类型、协议类型、访问权限等不同策略的访问控制表,就可以有效地限制广播组或共享域的大小。

(3) 简化了网络管理:一方面,可以不受网络用户的物理位置限制而根据用户需求进行网络管理,如同一项目或部门中的协作者,功能上有交叉的工作组,共享相同网络应用或软件的不同用户群。另一方面,由于 VLAN 可以在单独的交换设备或跨多个交换设备实现,也会大大减少在网络中增加、删除或移动用户时的管理开销。增加用户时只要将其所连接的交换机端口指定到其所属于的 VLAN 中即可;而在删除用户时只要将其 VLAN 配置撤销或删除即可;在用户移动时,只要他们还能连接到任何交换机的端口,则无须重新布线。

(4) 提供了基于第 2 层的通信优先级服务:在最新的以太网技术如千兆位以太网中,基于与 VLAN 相关的 IEEE 802.1q 标准可以在交换机上为不同的应用提供不同的服务(如传输优先级等)。

总之,VLAN 是交换式网络的灵魂,其不仅从逻辑上对网络用户和资源进行有效、灵活、简便管理提供了手段,同时提供了极高的网络扩展和移动性。但是请注意,尽管 VLAN 具有众多的优越性,但它并不是一种新型的局域网技术,而是一种基于现有交换机设备的网

络管理技术或方法，是提供给用户的一种服务。

5.6.4 虚拟局域网的工作方式

1. 基于交换端口的 VLAN

这种方式是把局域网交换机的某些端口的集合，作为 VLAN 的成员。这些集合有时只在单台局域网交换机上，有时则跨越多台局域网交换机。虚拟局域网的管理应用程序，根据交换机端口的标识 ID，将不同的端口分到对应的分组中，分配到一个 VLAN 的各个端口上的所有站点都在一个广播域中，它们相互之间可以通信，不同的 VLAN 站点之间进行通信需经过路由器来进行。这种 VLAN 方式的优点在于简单，容易实现，从一个端口发出的广播，直接发送到 VLAN 内的其他端口，也便于直接监控。它的缺点是自动化程度低，灵活性不好。比如，不能在给定的端口上支持一个以上的 VLAN；一个网络站点从一个端口移动到另一个新的端口时，如新端口与旧端口不属于同一个 VLAN，则用户必须对该站点重新进行网络地址配置。

2. 基于 MAC 地址的 VLAN

这种方式的 VLAN，要求交换机对站点的 MAC 地址和交换机端口进行跟踪，在新站点入网时，根据需要将其划归至某一个 VLAN。不论该站点在网络中怎样移动，由于其 MAC 地址保持不变，因此用户不需要对网络地址重新配置。然而所有的用户必须明确地分配给一个 VLAN，在这种初始化工作完成后，对用户的自动跟踪才成为可能。在一个大型网络中，要求网络管理人员将每个用户一一划分到某一个 VLAN 中，是十分烦琐的。

3. 基于路由的 VLAN

路由协议工作在 7 层协议的第 3 层——网络层，即基于 IP 和 IPX 的转发，它是利用网络层的业务属性来自动生成 VLAN，把使用不同路由协议的站点分在相对应的 VLAN 中。IP 子网 1 为第 1 个 VLAN，IP 子网 2 为第 2 个 VLAN，IPX 子网 1 为第 3 个 VLAN……以此类推。通过检查所有的广播和多点广播帧，交换机能自动生成 VLAN。这种方式构成的 VLAN，在不同的 LAN 网段上的站点可以属于同一 VLAN，同一物理端口上的站点也可分属于不同的 VLAN，从而保证了用户完全自由地进行增加、移动和修改等操作。这种根据网络上应用的网络协议和网络地址划分 VLAN 的方式，对于那些想针对具体应用和服务来组织用户的网络管理人员来说是十分有效的。它减少了人工参与配置 VLAN，使 VLAN 有更大的灵活性，比基于 MAC 地址的 VLAN 更容易做到自动化管理。

4. 基于策略的 VLAN

基于策略的 VLAN 的划分是一种比较灵活有效而直接的方式。这主要取决于在 VLAN 的划分中所采用的策略。目前常用的策略如下：

（1）按 MAC 地址划分；

（2）按 IP 地址划分；

（3）按以太网协议类型划分；

（4）按网络的应用划分。

5.6.5　虚拟局域网的实现

从实现的方式上看，所有 VLAN 均是通过交换机软件实现的。从实现的机制或策略划分，VLAN 分为静态 VLAN 和动态 VLAN。

VLAN 的实现方式包括静态 VLAN 和动态 VLAN 两种。

1. 静态 VLAN

在静态 VLAN 中，由网络管理员根据交换机端口进行静态的 VALN 分配，当在交换机上将其某一个端口分配给一个 VLAN 时，将一直保持不变直到网络管理员改变这种配置，所以又被称为基于端口的 VLAN。基于端口 VLAN 配置简单，网络的可监控性强。但缺乏足够的灵活性，当用户在网络中的位置发生变化时，必须由网络管理员将交换机端口重新进行配置。所以静态 VLAN 比较适合用户或设备位置相对稳定的网络环境。

2. 动态 VLAN

动态 VLAN 是指交换机上以联网用户的 MAC 地址、逻辑地址（如 IP 地址）或数据报协议等信息为基础将交换机端口动态分配给 VLAN 的方式。当用户的主机连入交换机端口时，交换机通过检查 VLAN 管理数据库中相应的关于 MAC 地址、逻辑地址（如 IP 地址）或数据报协议的表项，以相应的数据库表项内容动态地配置相应的交换机端口。以基于 MAC 地址的动态 VLAN 为例，网络管理员首先需要在 VLAN 策略服务器上配置一个关于 MAC 地址与 VLAN 划分映射关系的数据库，当交换机初始化时将从 VLAN 策略服务器上下载关于 MAC 地址与 VLAN 划分关系的数据库文件，此时，若有一台主机连接到交换机的某个端口时，交换机将会检测该主机的 MAC 地址信息，然后查找 VLAN 管理数据库中的 MAC 地址表项，用相应的 VLAN 配置内容来配置这个端口。这种机制的好处在于只要用户的应用性质不变，并且其所使用的主机不变（严格地说，是使用的网卡不变），则用户在网络中移动时，并不需要对网络进行额外配置或管理。但是，在使用 VLAN 管理软件建立 VLAN 管理数据库和维护该数据库时需要做大量的管理工作。总之，不管以何种机制实现，分配给同一个 VLAN 的所有主机共享一个广播域，而分配给不同 VLAN 的主机将不会共享广播域。也就是说，只有位于同一 VLAN 中的主机才能直接相互通信，而位于不同 VLAN 中的主机之间是不能直接相互通信的。

5.6.6　VLAN 间的互联方法

1. 传统路由器方法

所谓传统路由器方法，就是使用路由器将位于不同 VLAN 的交换端口连接起来，这种方法的缺点是：对路由器的性能有较高要求；同时如果路由器发生故障，则 VLAN 之间就不能通信。

2. 采用路由交换机

如果交换机本身带有路由功能，则 VLAN 之间的互联就可在交换机内部实现，即采用第 3 层交换技术。第 3 层交换技术也叫路由交换技术，是各网络厂家最新推出的一种局域网技术，具有良好的发展前景。它将交换技术（Switching）和路由技术（Routing）相结合，很

好地解决了在大型局域网中以前难以解决的一些问题。

5.7 网络操作系统

5.7.1 网络操作系统概述

1. 网络操作系统的定义和功能

网络操作系统(Network Operation System,NOS)是指能使网络上多台计算机方便而有效地共享网络资源,为用户提供所需的各种服务的操作系统软件。

为实现有效的资源共享,首先要提供网络通信功能或协议的支持,另外还要提供资源共享的途径及解决多个用户对资源需求冲突的能力。所以网络操作系统除了具备单机操作系统所需的功能(如内存管理、CPU 管理、输入输出管理、文件管理等)以外,还应具备如下一些网络控制、管理和服务功能。

(1) 提供高效可靠的网络通信能力,如对网络协议、网络硬件的支持。例如,在 Windows 2000/2003 操作系统中,就有对 TCP/IP、NetBEUI、DLC 等多种协议的支持,同时还提供了多种网络硬件的驱动程序。

(2) 提供多项网络服务功能,如远程作业输入及处理服务功能、文件传输服务功能、电子邮件服务功能、远程打印服务功能等。我们经常听说的 Telnet、FTP、E-mail 等就是该类服务功能的典型例子。

(3) 提供网络资源管理、系统管理功能,如文件系统管理、网络服务进程的建立和管理、网络活动的监控和网络测试工具等。Windows 2000/2003 中的事件查看器就提供对一些网络安全方面的问题进行监视的功能。

(4) 提供对网络用户的管理。几乎所有的操作系统都提供了用户管理功能,用户管理功能所提供的用户访问控制机制有效地管理和控制了用户对网络资源的访问。用户必须提供合法的用户账号并在授权范围内访问网络资源就是用户管理的具体体现。

2. 网络操作系统的组成

NOS 通常有两个基本的组成部分,即运行在服务器上的操作系统和运行在每台 PC 或桌面工作站上的客户端操作系统软件。服务器操作系统的主要功能是控制服务器的操作,管理存储在服务器上的文件,提供对用户的集中管理,支持多用户和多任务的工作环境以解决多个用户对资源需求时的冲突。客户端操作系统的主要功能是提供客户访问网络及网络资源的能力,而这些网络资源通常由网络服务器提供。

5.7.2 常见的网络操作系统

网络操作系统是网络设计与实施过程中要考虑的关键因素之一。目前,可供选择的网络操作系统多种多样,常见的有 Windows、UNIX、Linux、NetWare 等。

1. Windows 操作系统

Windows 网络操作系统是一个产品系列。Microsoft 公司在 1993 年推出第一代网络

操作系统产品 Windows NT 3.1，随着 Windows NT 3.1 问世，Microsoft 正式加入网络操作系统的市场角逐。时至今日，微软公司先后对 Windows 网络操作系统不断进行改进，陆续推出 Windows NT 3.5、Windows NT 4.0、Windows Server 2000 家族以及现在的 Windows Server 2003。Windows 系列网络操作系统的主要特点有以下几个方面。

(1) 可靠性。衡量一个网络操作系统的可靠性不是一朝一夕的事，无论 Microsoft 在软件界的地位多高，它新推出的 Windows NT(2000/2003)在未经历相当时间的检验之前，系统的可靠性、稳定性还是未知数，慎重的客户也不会盲目地一下子拥向 Windows NT，所以现在比较慎重的用户还是在坚持应用 UNIX。

(2) 新概念和新技术。首先，因为 Windows NT 是最新设计的网络操作系统，它自然而然就会采用最新的概念和最新的技术。以前的网络操作系统在设计时根本不会考虑到的因素，Windows NT 的设计者都考虑到了，这绝不是说别的系统不够先进或没有远见，只是受历史条件和当时的技术发展因素所限，不可能预见。

(3) 友好的界面。Windows NT 具有友好的界面。统一的界面风格是 Windows 系列开拓市场的强有力的武器。简单的操作使用户免予记诵繁杂的命令而一上手就可以使用，并且更重要的是，Windows NT 提供的功能以及开发工具绝不逊色于任何别的优秀系统。

(4) 丰富的配套应用。Microsoft 公司在软件界有着特殊的地位，一方面它是平台提供商；另一方面它也是应用提供商。这样的双重身份使得 Microsoft 的产品具有一些特别之处。对于网络操作系统产品而言，因为 Microsoft 本身就是应用提供商，所以在其上的应用服务就不会匮乏。而且，因为是出自同一公司之手，因而应用和平台的结合应当是优秀的。应用可以充分利用 Microsoft 的平台优势，平台也能充分支持其开发的应用。此外，新出的 Windows 2000/2003 的 VLM 将提供大内存寻址能力和动态目录服务，弥补了 Microsoft 在这方面的一个不足。此外，Microsoft 的“零管理”将大大降低系统的管理成本。

正是上述优越的性能，使得 Microsoft 的 Windows 网络操作系统系列产品后来居上，在当今的网络操作系统市场占有举足轻重的地位。

2. UNIX 操作系统

UNIX 最早是指由美国贝尔实验室发明的一种多用户、多任务的通用操作系统。经过长期的发展和完善，目前已成长为一种主流的操作系统技术和基于这种技术的产品大家族。其中最为著名的有 SCO XENIX、SNOS、Berkeley BSD、AT&T 系统 V。由于 UNIX 具有技术成熟、可靠性高、网络和数据库功能强、伸缩性突出和开放性好等特色，可满足各行各业的实际需要，特别能满足企业重要业务的需要，已经成为主要的工作站平台和重要的企业操作平台。目前，每年仍以两位数字以上的速度稳步增长。早期 UNIX 的主要特色是结构简练、便于移植和功能相对强大。经过多年的发展和进化，又形成了一些极为重要的特色，其中主要包括以下几点。

(1) 技术成熟，可靠性高：经过 30 年开放式道路的发展，UNIX 的一些基本技术已变得十分成熟，有的已成为各类操作系统的常用技术。实践表明，UNIX 是能达到主机(Mainframe)可靠性要求的少数操作系统之一。目前，许多 UNIX 主机和服务器在国内外的大型企业中每天 24 小时、每年 365 天不间断地运行。

(2) 极强的伸缩性(Scalability)：UNIX 是世界上唯一能在笔记本电脑、PC、工作站直至巨型机上运行的操作系统，而且能在所有主要体系结构上运行。迄今为止，世界上没有第

二个操作系统能做到这一点。此外，由于 UNIX 操作系统能很好支持 SMP、MPP 和 Cluster 等技术，使其可伸缩性又有了很大的增强。

(3) 强大的网络功能：网络功能强是 UNIX 操作系统的又一重要特色，作为 Internet 技术基础和异种机连接重要手段的 TCP/IP 协议就是在 UNIX 上开发和发展起来的。TCP/IP 是所有 UNIX 操作系统不可分割的组成部分。因此，UNIX 服务器在 Internet 服务器中占 70% 以上，占绝对优势。此外，UNIX 还支持所有常用的网络通信协议，包括 NFS、DCE、IPX/SPX、SLIP、PPP 等，使得 UNIX 操作系统能方便地与已有的主机系统，以及各种广域网和局域网相连接，这也是 UNIX 具有出色的互操作性(Interoperability)的根本原因。

(4) 强大的数据库支持能力：由于 UNIX 具有强大的数据库支持能力和良好的开发环境，多年来，所有主要数据库厂商，包括 Oracle、Informix、Sybase、Progress 等，都把 UNIX 作为主要的数据库开发和运行平台，并创造出一个又一个性价比的新纪录。

(5) 功能强大的开发平台：UNIX 操作系统从一开始就为软件开发人员提供了丰富的开发工具，成为工程工作站的首选和主要的操作系统与开发环境。可以说，工程工作站的出现和成长与 UNIX 是分不开的。迄今为止，UNIX 工作站仍是软件开发厂商和工程研究设计部门的主要工作平台。有重大意义的软件新技术几乎都出现在 UNIX 上，如 TCP/IP、WWW 等。

(6) 开放性好：开放性是 UNIX 最重要的本质特征。开放系统概念的形成与 UNIX 是密不可分的。UNIX 是开放系统的先驱和代表。由于开放系统深入人心，几乎所有厂商都宣称自己的产品是开放系统，确实每一种系统都能满足某种开放的特性，如可移植性、兼容性、伸缩性、互操作性等。但所有这些系统与开放系统的本质特征——不受某些厂商的垄断和控制相去甚远，只有 UNIX 完全符合这一条件。

3. Linux 操作系统

Linux 是一个免费的、提供源代码的操作系统。Linux 脱胎于 UNIX，所以其很多性能和特点与 UNIX 极其相似。Linux 最早出现在 1992 年，由芬兰赫尔辛基大学的一个大学生 Linus B. Torvalds 首创，后来在全世界各地由成千上万的 Internet 上的自由软件开发者协同开发，不断完善。经过十多年的发展，它已完全进入了成熟阶段，越来越多的人认识到它的价值，从 Internet 服务器到用户的桌面，从图形工作站到 PDA 的各种领域都在广泛使用。Linux 下有大量的免费应用软件，从系统工具、开发工具、网络应用，到休闲娱乐、游戏等。更重要的是，它是目前安装在个人计算机上的最可靠、最强壮的操作系统。

Linux 作为一个置于共用许可证(General Public License，GPL)保护下的自由软件，任何人都可以免费从分布在全世界各地的网站上下载。目前，Linux 的发行版本种类很多，最主要的几个发行版本为 Red Hat Linux、Slackware、Debian Linux、S. u. S. e Linux 等，最近国内也有人搞了自己的发行版本，如联想公司的幸福 Linux 以及冲浪平台的 Xteam Linux。

4. NetWare 操作系统

Novell 公司的 NetWare 网络操作系统是目前世界上应用非常广泛的微机局域网操作系统之一。NetWare V6 是 Novell 公司的最新产品，也是目前局域网里非常优秀的网络操作系统之一。NetWare 推出时间比较早，经过多年的发展，已可以提供非常稳定的运行性能。在一个 NetWare 网络中允许有多台服务器，并可采用一般的 PC 担当服务器。

NetWare 的主要优点如下。

(1) 强大的文件及打印服务能力。NetWare 以其强大的文件及打印服务能力而久负盛名。NetWare 能够通过文件及目录高速缓存,将那些读取频率较高的数据预先读入内存,来实现高速文件处理;在 NetWare 中,还可以将打印服务软件装入像文件服务器这样的硬件当中,以方便地实现打印机资源共享。

(2) 兼容性及系统容错能力。较高版本的 NetWare(如 NetWare 3.X)不仅能与不同类型的计算机兼容,而且还能与不同类型的操作系统兼容。另外,它所具备的 SFT(系统差错容限)和 TTS(事务跟踪系统)技术,能够在系统出错时及时进行自我修复,大大降低了因重要文件和数据丢失所带来的不必要的损失。

(3) 比较完备的安全措施。NetWare 对入网用户进行注册登记,并采用四级安全控制原则以管理不同级别的用户对网络资源的使用。在 NetWare 4.1 中,还采用了名为 NDS (Net Directory Service,网络目录服务)的技术,使用户无须了解打印机或文件位于哪台服务器中,就能使用该打印机或文件。NetWare 主要的不足之处是工作站资源无法直接共享,安装及管理维护较为复杂,多用户同时获取文件及数据时会导致网络效率降低,以及服务器的运算功能没有得到充分发挥等缺点。

课后习题

1. 术语解释

介质访问控制　以太网　快速以太网　WLAN　VLAN　MAC　LLC　NOS

2. IEEE 802.3 标准是(　　)。

A. 逻辑链路控制

B. CSMA/CD 访问方法和物理层规范

C. 令牌总线访问方法和物理层规范

D. 令牌环网访问方法和物理层规范

3. 无线局域网所采用的协议为(　　)。

A. CSMA/CD　　B. Taken Ring　　C. CSMA/CA　　D. PPP

4. 以太网媒体访问控制技术 CSMA/CD 的机制是(　　)。

A. 争用带宽　　B. 预约带宽

C. 循环使用带宽　　D. 按优先级分配带宽

5. IEEE 802.3 的物理层协议 10Base-T 规定从网卡到集线器的最大距离为(　　)m。

A. 100　　B. 185　　C. 500　　D. 850

6. 在局域网中,MAC 指的是(　　)。

A. 逻辑链路控制子层　　B. 介质访问控制子层

C. 物理层　　D. 数据链路层

7. MAC 地址又称网卡地址,是用于在物理上标识主机的,以下四个选项中,只有选项(　　)所表示的 MAC 地址是正确的。

A. A2-16　　B. 00-02-60-07-A1-1C

C. 202.168.1.32　　　　D. 01-02-6G-70-A1-EC

8. 令牌总线的媒体访问控制方法是由(　　)定义的。

A. IEEE 802.2　　B. IEEE 802.3　　C. IEEE 802.4　　D. IEEE 802.5

9. 在计算机网络中,所有的计算机均连接到一条通信传输线路上,在线路两端连有防止信号反射的装置,这种连接结构被称为(　　)。

A. 总线型结构　　B. 环形结构　　C. 星形结构　　D. 网状结构

10. 令牌环网中某个站点能发送帧是因为(　　)。

A. 最先提出申请　　B. 优先级最高　　C. 令牌到达　　D. 可随机发送

11. IEEE 802 局域网参考模型与 OSI 参考模型有何异同之处?

12. 试说明 10Base-5、10Base-2、10Base-T、10Base-F、1Base-5、10Broad-36 所代表的意思。

13. 假定 1km 长的 CSMA/CD 网络的数据传输速率为 1Gbps,设信号在网络上的传输速率为 200000km/s,求能够使用此协议的最短帧长。

14. 以太网使用的 CSMA/CD 协议是以争用方式接入共享信道,这与传统的时分多路复用(TDM)相比优缺点如何?

15. 以太网交换机有何特点?用它怎样组成虚拟局域网?

16. 常用的介质访问方法有哪些?

17. 局域网有哪几种常见的拓扑结构?各有何特点?

18. 与有线局域网比较,无线局域网具有哪些优越性?

19. 什么是 VLAN?引入 VLAN 有哪些优越性?VLAN 是如何实现的?

20. 试对 Windows、Linux、UNIX、NetWare 4 种操作系统的特点作一比较。

第6章 网络层

学习目的

本章在前面物理层与数据链路层学习的基础上，系统地介绍网络层的基本概念、网络层向传输层提供的服务功能、IP 协议、子网与子网掩码、流量控制算法、路由与路由选择基本算法等内容。

本章内容主要围绕着 IP 协议展开，通过本章的学习为进一步研究 Internet 工作原理与实现打下坚实的基础。

学习要求

理解：网络层与网络互联的基本概念。

掌握：IP 地址的基本概念与分类方法。

掌握：路由与路由选择协议的概念。

掌握：IP 协议的基本内容。

掌握：地址解析的基本概念和方法。

掌握：子网与子网掩码、无分类编址的概念。

掌握：拥塞控制的概念与其算法。

了解：网际协议 IPv6 及移动 IP。

在前面的章节中，已经学习了物理层和数据链路层的功能及其实现，并对典型局域网有了一定了解。各种网络技术的主要区别就在于物理层和数据链路层，那么当这些不同的网络被互联在一起时，就会出现异构网络互联的问题。而且当互联网的规模增加时，相互通信的源和目的节点之间可能会存在一系列的中间节点，从而还会带来路径选择的问题。这些问题都有待于 OSI 参考模型的第 3 层即网络层去解决。

6.1 网络层功能概述

6.1.1 网络层的必要性

数据链路层能利用物理层所提供的比特流传输服务实现相邻节点之间的可靠数据传输，也就是说，数据链路层只能将数据帧由传输介质的一端送到另一端。网络层是 OSI 参考模型中的第 3 层，它建立在数据链路层的两个相邻端点之间的数据帧的传送功能之上，将数据从源端经过若干中间节点传送到目的端，从而向传输层提供最基本端到端的数据传送服务。

以图 6.1 为例，源主机 DTE1 和中间节点 DCE1 为相邻节点，DCE1 分别与中间节点 DCE2、DCE3 和 DCE4 为相邻节点，DCE2 的相邻节点则包括了 DCE3 和 DCE6，数据链路

层已经有效解决了诸如这些相邻节点之间的数据传输问题。但是,从图 6.1 中可以看出,如果要完成从源节点 DTE1 到目标节点 DTE2 的数据传输,不可避免地要历经一些中间节点,这些中间节点构成了从源到目标的多条网络路径,从而导致了路径选择问题。例如,当 DCE1 收到从 DTE1 来的数据后,它马上面临着是从 DCE2 还是 DCE3 或者是 DCE4 进行数据转发的问题,而数据链路层并没有提供这种实现从源到目标的数据传输所必需的路径选择功能。

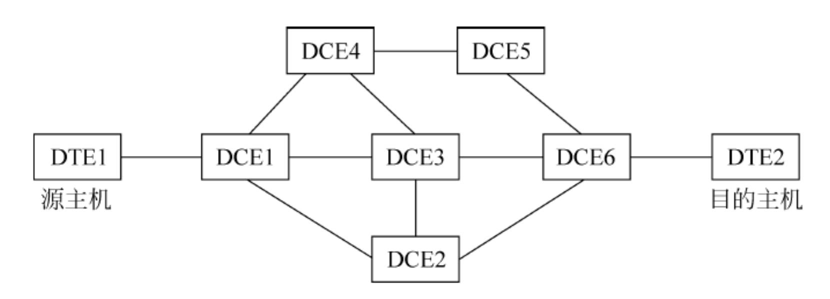


图 6.1 网络中间节点和网络路径的示例

也许有的同学会提出问题,既然数据链路层能够以物理地址来标识网络中的每一个节点,为什么不能绕开路径选择问题而直接利用物理层地址实现主机寻址呢?请同学们注意,物理地址是一种平面化的地址,其地址编码是按照顺序分配方式获得的,在地址编码中不包含任何有关主机所在网络的结构信息。利用这种非结构化的地址信息在规模较大的网络中实现主机寻址几乎是不可能的。以以太网为例,当源主机和目的主机分别位于同一个网桥或交换机端口所直接相连的物理网段时,这种物理寻址方式的确可以比较方便地定位到目的主机。但是,若目的主机不在同一个网桥或交换机直接所连物理网段时,网桥或交换机就只能通过向所有与之相连的其他网桥或交换机进行洪泛的方式来间接地找到目标节点。从表面上看,这种方式似乎也是可行的,但当网络互联规模增大时,会因为大量的洪泛流量而导致网络性能下降甚至是瘫痪。也就是说,通过物理地址直接寻址的方式只能适用于规模非常小的网络环境,在绝大多数情况下必须借助网络层的路径选择功能来实现网络中的逻辑寻址。

6.1.2 网络层的功能

网络层涉及将源主机发出的分组经由各种网络路径到达目的主机,其利用了数据链路层所提供相邻节点之间的数据传输服务,向传输层提供了从源到目的的数据传输服务。网络层是处理端到端(End to End)数据传输的最低层,但同时又是通信子网的最高层。如图 6.2 所示,资源子网中的主机具备了 OSI 参考模型中所有 7 层的功能,但通信子网中的主机因为只涉及通信问题而只拥有 OSI 参考模型的第 3 层。所以网络层被看成是通信子网与资源子网的接口,即通信子网的边界。

为了有效地实现源主机到目的主机的分组传输服务,网络层需要提供多方面的具体功能。

首先,需要规定该层协议数据单元的类型和格式。网络层的协议数据单元被称为分组(Packet)。与其他各层的协议数据单元类似,分组是网络层协议功能的集中体现,其中要包

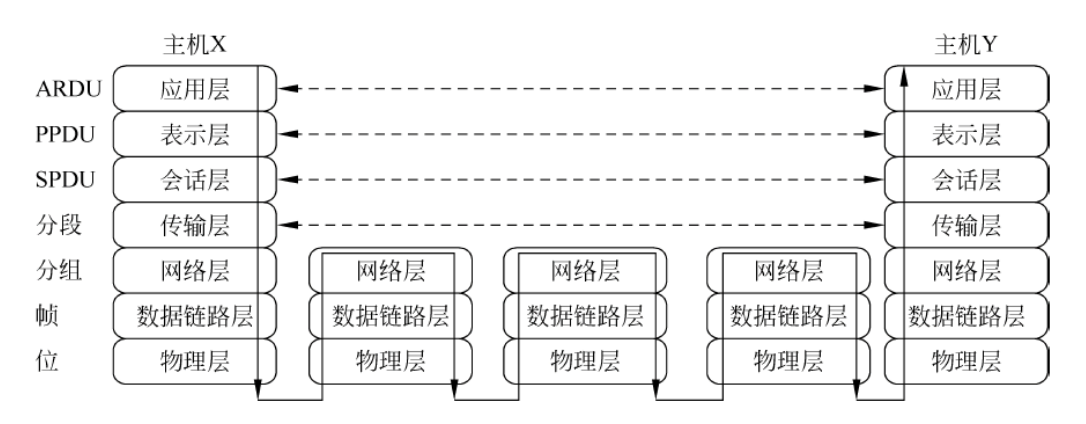


图 6.2　网络层的地位与作用

括实现该层功能所必需的控制信息如收发双方的网络地址等。

其次，要了解通信子网的拓扑结构，并通过一定的路由算法为分组实现进行最佳路径选择，最佳路径选择又被称为路由(Routing)。

再次，在为分组选择路径时还要注意既不要使某些路径或通信线路处于超负载状态，也不能让另一些路径或通信线路处于空闲状态，即所谓的拥塞控制和负载平衡。通常，当网络负载过重、带宽不够或通信子网中的路由设备性能不足时都可能导致拥塞。

最后，在从源主机到目的主机所经历的网络属于不同类型时，网络层还要协调好不同网络间的差异即所谓解决异构网络互联的问题。所谓异构，是指网络技术、通信协议、计算机体系结构或操作系统上的差异性。具体表现如下。

(1) 网络类型的不同：广域网、城域网或局域网。

(2) 网络实现技术的不同：同一种类型的网络可以用不同的物理层和数据链路层技术去实现，如局域网中的以太网、FDDI、令牌环网和 IEEE 802.11 无线局域网都有各自的物理层和 MAC 子层的实现方式。

(3) 使用不同类型的通信协议：如 HDLC 协议、PPP 协议或 SDH 等。

(4) 不同类型的计算机系统：计算机的机型包括大型机、小型机、PC 和工作站等，不同的机型可能采用不同的计算机体系结构。

(5) 所使用的操作系统不同：如 UNIX、Linux、Windows、NetWare 和 OS2 等。

通常，当网络覆盖范围增大、网络互联程度增加时，网络的异构性程度也会随之增加，网络层必须设法解决异构网络互联的问题，以满足用户在扩大网络覆盖范围、增强网络互联性上的需求。

根据分层的原则，网络层在为传输层提供分组传输服务还必须做到：服务与通信子网技术无关，即通信子网的数量、拓扑结构及类型对于传输层是透明的；传输层所能获得的地址应采用统一的方式，以使其能跨越不同的 LAN、MAN 和 WAN 而提供互联网中的寻址能力。上述这些要求也就是网络层设计的基本目标。

6.1.3　网络层所提供的服务

网络层提供给传输层的服务有面向连接和面向无连接之分。所谓面向连接，是指在数据传输之前双方需要为此建立一种连接，然后在该连接上实现有次序的分组传输，直到数据

传输完毕才释放连接；面向无连接则不需要为数据传输事先建立连接，其只提供简单的源和目的之间的数据发送与接收功能。网络层服务方式的不同主要取决于通信子网的内部结构。面向无连接服务在通信子网内通常以数据报(Datagram)方式实现。在数据报服务中，每个分组都必须提供关于源和目的的完整地址信息，通信子网根据地址信息为每一个分组独立进行路径选择。数据报方式的分组传输可能会出现丢失、重复或乱序的现象。面向连接服务则通常采用虚电路(Virtual Circuit，VC)方式实现。虚电路是指通信子网为实现面向连接服务而在源与目的之间所建立的逻辑通信链路。

1. 数据报

在数据报的操作方式中，每个分组在传输过程中是单独处理的。每个分组称为一个数据报，它是一个完备、独立的数据实体，每个数据报自身携带从源计算机传递到目的计算机的信息。一个节点收到一个数据报后，根据数据报中的地址信息和节点所存储的路由信息，找出一个合适的路径，把数据报发送到下一个节点。在数据报操作方式中，每个数据报自身携带有足够的信息，它的传送是被单独处理的。整个数据报传送过程中，不需要建立虚电路，网络节点为每个数据报进行路由选择，因此不能保证各数据报按顺序到达目的节点，有些还可能会丢失。整个过程中，没有虚电路建立，但是要为每个数据报进行路由选择。其特点如下。

(1) 同一报文的不同分组可以由不同的传输路径通过通信子网。

(2) 同一报文的不同分组到达目的节点时可能出现乱序、重复与丢失现象。

(3) 每一个分组在传输过程中都必须带有目的地址与源地址。

(4) 数据报方式的报文传输延迟较大，适用于突发性通信，不适用于长报文、会话式通信。

2. 虚电路

虚电路又称为虚连接或虚通道，是分组交换的两种传输方式中的一种。在通信和网络中，虚电路是由分组交换通信所提供面向连接的通信服务。在两个节点或应用进程之间建立起一个逻辑上的连接或虚电路，就可以在两个节点之间依次发送每一个分组，接收端收到分组的顺序必然与发送端的发送顺序一致，因此接收端无须负责在收集分组后重新进行排序。虚电路协议向高层协议隐藏了将数据分割成段、包或帧的过程。

表 6.1 所示为数据报与虚电路的差别比较。

表 6.1 数据报与虚电路比较

分类 比较	数 据 报	虚 电 路
延时	分组传输延时	电路建立，分组传输延时
路由选择	每个分组单独选择路由	建立虚电路时选择路由，以后所有分组都使用该路由
状态信息	子网无须保存状态信息	每个节点要保存一张虚电路表
地址	每个分组需要完整的源和目的地址	每个分组只需包含一个虚电路号
传输质量	同一报文的不同分组会出现乱序、重复或丢失	同一报文的不同分组不会出现乱序、重复或丢失
连接设置	不需要	需要
通信效率	相对高	相对低
协议复杂度	相对低	相对高

6.2　数据交换方式

最简单的数据通信形式是两个站点直接使用物理线路进行通信，但如果两个站点相距遥远或者要进行多站点之间的通信，采用直接连接显然是不合适的。因为任意两个站点直接专线连接费用昂贵，例如 n 个站点全连通，即其中任一站点同其他所有站点($n-1$)个都有专线相连，则总共需要 $n(n-1)/2$ 条专线，这显然是不经济的。解决这一问题的方法就是采用交换技术，所谓交换技术是采用交换机或节点机等交换系统，通过路由选择技术在进行通信的双方之间建立物理的或逻辑的连接，形成一条通路，实现通信双方的信息传输和交换的一种技术。

常用的数据交换方式可分为两大类：电路交换方式(Circuit Switching)和存储转发交换方式(Store and Forward Switching)。存储转发交换方式按照被转接的信息单位不同，又可分为报文交换和报文分组交换。下面分别介绍这几种交换方法。

6.2.1　电路交换

在电路交换网络中，通过网络节点在两个工作站之间建立一条专用的通信电路。最普通的电路交换例子是公用电话交换网(PSTN)。使用电路交换方式进行通信时，两个工作站之间使用实际的物理连接或逻辑连接，这种连接由节点的各段电路组成，每一段电路都为该连接提供一条通道。电路交换方式的通信过程包括以下 3 个阶段。

1. 电路建立

在传输任何数据之前，都必须建立端到端(站到站)的线路，即在源节点和目的节点间建立一条由各个中间交换节点的分段连接所组成的通信电路。如图 6.3 所示的网络，站点 A 向节点 1 发出请求，要求与 B 站通信。由于站点 A 到节点 1 以及站点 B 到节点 6 均只有专用线路，所以节点 1 必须接通一条到节点 6 的电路，节点 1 到节点 6 的电路可以有多种选择：比如 1→2→6，1→4→3→6 等。假设根据路由选择的规则选择了电路 1→4→3→6，那么就建立了从站点 A 到站点 B 的电路 A→1→4→3→6→B。

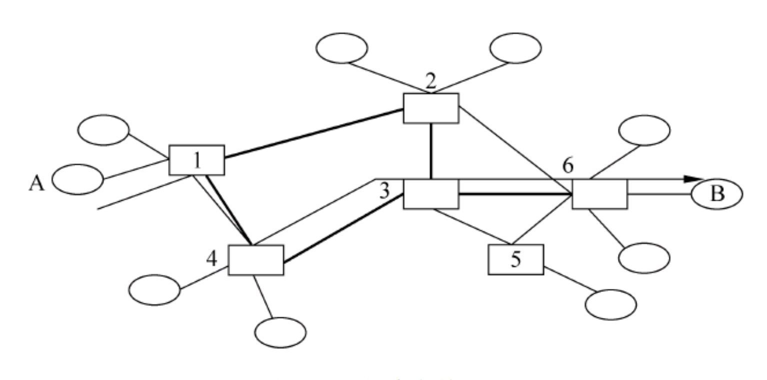


图 6.3　电路交换过程

2. 数据传输

一旦通信电路建立起来，就可通过这条专用电路从站点 A 通过网络传输数据到站点

B。其中传输的数据可以是数字数据,也可以是模拟数据。

3. 电路拆除

数据通信结束后,应拆除电路,供其他用户使用。通常是由两个站点中的一个站点来完成这一动作。拆除线路信号必须传输到电路所经过的各个节点,以便重新分配资源。

电路交换方式,在传输数据之前建立连接,有延迟;在电路建立后就专用该电路,即使没有数据传输也要占用电路,所以利用率较低。然而,一旦建立了连接,网络对于用户实际上是透明的;用户可以以固定的速率传输数据,除了传输延迟外,不再有其他的延迟。电路交换能适应实时性传输,但如果通信量不均匀,容易引起拥塞。

6.2.2 报文交换

报文交换(Message Switching)属于存储交换,它不需要在两个站之间建立一条专用通路。存储交换的主要原理是:把待传输的信息存储起来,等到信道空闲时发出去。只要存储时间足够长,就能够把信道忙碌和空闲的状态均匀化,大大压缩了必需的信道容量和转接设备容量。但是,这种方式对于有实时性要求的信息传输是不允许的,而对于数据通信则是合适的。存储交换具有存储信息的能力,所以能平滑通信和充分利用信道。

报文交换工作过程是:发信端将发往收信端(目的地)的信息分割成一份份的报文正文,连同收信地址等辅助信息形成一份份的报文,首先发往本地的交换中心(或交换局),然后由交换中心将每份报文完整地存储起来;由于报文一般较长,往往将它存入联机的大容量存储器或脱机的大容量存储器中,当等到去目的地的线路空闲时,再将一份份报文转发到下一个交换中心,然后再转到目的地。目的地收信交换中心将收到的各份报文按原来的顺序进行装配,而后将完整的信息交付给目的地收信的计算机或终端设备,如图 6.4 所示。报文从站点 A 出发经过节点 1、节点 2 和节点 6 的存储转发,最后到达站点 B。

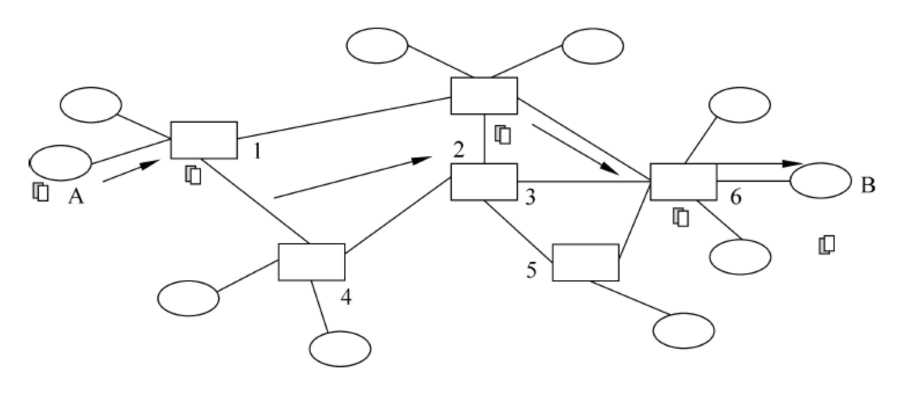


图 6.4 报文交换过程

报文交换方式以报文为单位发送信息。单个报文包括 3 部分内容:报头(Header)、报文正文(Message)和报尾(Trailer)。报头由发信端地址、终点收信端地址及其他辅助信息组成。有时也省去报尾,但此情况下的单个报文必须有统一的固定长度。报文交换方式没有拨号呼叫,由报文的报头始终控制其到达目的地。

由于交换中心有存储能力和信息处理能力,因而在信息传递过程中可以通过交换机进行速率变换、符号变换及格式变换等,使得不同类型的终端设备可以相互连接,同一报文也

可由交换中心按需要转发到几个收信地；多条低速线路可以集中化、高速化，从而提高线路的利用率。此外，以报文为单位来占用信道，可复用线路(即多个用户共用一条信道)，也可使终端用户在思考问题、等待应答的各种无效时间内，不再独占信道，进一步提高了线路的利用率。

报文交换和电路交换相比有以下优点。

(1) 电路效率较高，因为许多报文可以分时共享一条节点到节点的通道。

(2) 不需要同时使用发送器和接收器来传输数据，网络可以在接收器可用之前，暂时存储这个报文。

(3) 在电路交换网上，当通信量变得很大时，就不能接收某些呼叫。而在报文交换网上，却仍然可以接收到报文，虽然报文被缓存导致传输延迟增加，但不会引起拥塞。

(4) 报文交换系统可以把一个报文发送到多个目的地。

(5) 根据报文的长短或其他特征能够建立报文的优先权，使得一些短的、重要的报文优先传递。

(6) 报文交换网可以进行速度和代码的转换，因为每个站都可以用它特有的数据传输率连接到其他节点，所以两个不同传输速率的站点之间也可以连接。

但报文交换网不能满足实时性或交互性的通信要求，经过网络的延迟时间较长，而且由于负载不同，延迟时间有较大的变化。这种方式不能用于声音连接，也不适合交互式终端到计算机的连接。

6.2.3　报文分组交换

报文分组交换(Packet Switching)方式是1964年提出来的，简称为分组交换或包交换，最早在ARPANET上得以应用，它试图兼有报文交换和电路交换的优点，而使两者的缺点最少。报文分组交换方式与报文交换方式相比，它采用了较短的格式化的信息单位，称为报文分组，简称报文组(Packet)。在报文分组交换网中，典型数据单位分组的长度限制在一千比特到数千比特；而在报文交换网中，报文长度远比分组长得多。

ITU给报文分组下的定义是：一组包含数据和呼叫控制信号(例如地址)的二进制数，对它作为一个组合整体加以交换，这些数据、呼叫信号以及可能附加的差错控制信息是按规定的格式排列的。由于它在发送端将报文分割成更小的报文分组，使它适合在交换机(计算机)的主存储器中存储转发，所以比起报文交换方式，报文分组交换能改善传输的接续时间和传输延迟时间。

图6.5表示了报文分组交换过程：当报文从站点A出发到达节点1后，分成多个分组，每个分组各自选择不同的路径，最后都到达节点6后，重新装配成报文，传输给站点B。

由于采用分组传输以后，发送信息时需要把报文信息拆卸并加入分组报头，即将报文转换成分组信号；接收时还需要去掉分组报头，将分组数据装配成报文信息。所以，用于控制和处理数据传输的软件较复杂，同时对通信设备的要求也较高。以这种方式构成的通信网可以采用分布式控制的自适应路由选择技术，即根据通信量当前的通路情况(通路故障、交换机故障均可动态地得到反映)及通信量情况，选择最佳的路由(例如，以报文分组传输延迟时间最小为最佳依据)，以便网络中各信道的流量趋于平衡。报文分组交换采用两种方法来

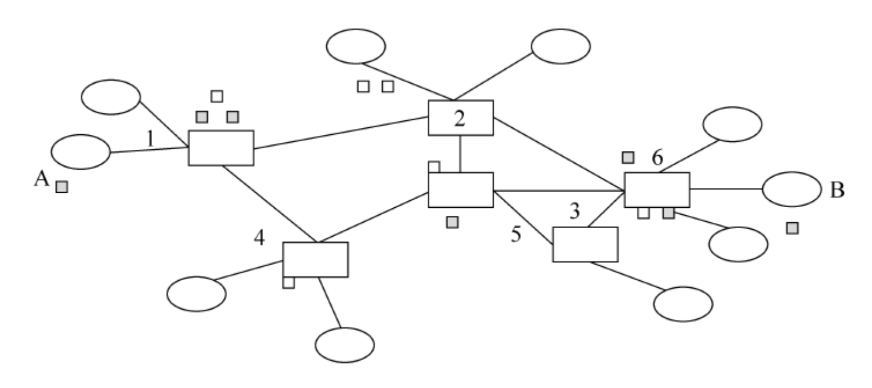


图 6.5　报文分组交换过程

管理分组流：数据报和虚电路。

6.2.4　三种交换技术比较

图 6.6 有助于了解三种交换技术的有关性能，但是实际的性能取决于诸多因素，其中包括：站的数目、节点的数目和排列、系统的总负载、两个站之间典型的交换长度（时间长度和数据长度）。

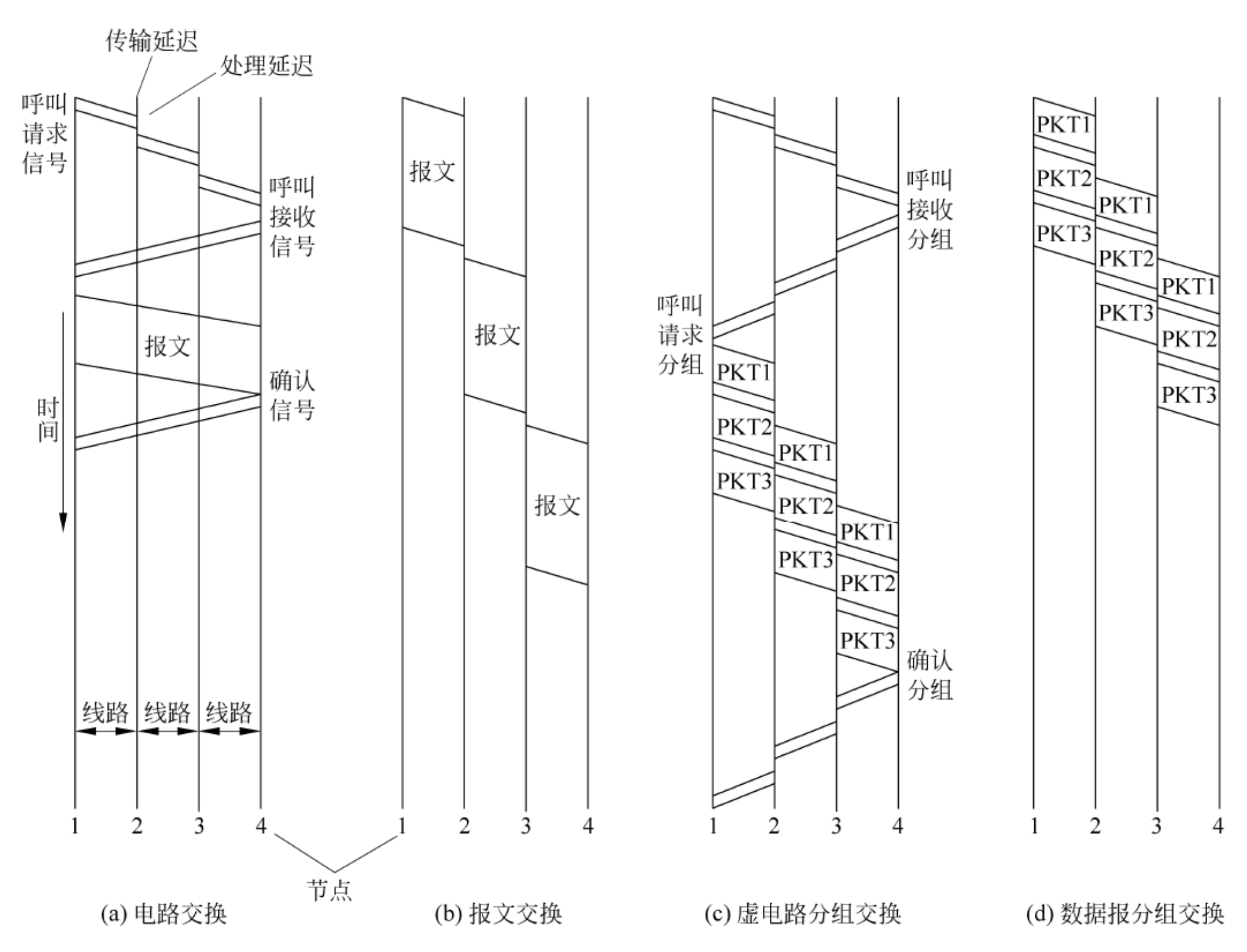


图 6.6　各种通信技术的工作时序

不同的交换技术适用于不同的场合,说明如下。

(1) 对于交互式通信来说,报文交换是不合适的。

(2) 对于较轻的或间歇式负载来说,电路交换是最合算的,因为可以通过电话拨号来使用公用电话系统。

(3) 对于两个站之间很重的和持续的负载来说,使用租用的电路交换是最合算的。

(4) 当有一批中等数量数据必须交换到大量的数据设备时,可用报文分组交换方法,这种技术的线路利用率是最高的。

(5) 数据报分组交换适用于短报文,能具有灵活性的报文。

(6) 虚电路分组交换适用于长交换,能减轻各站点的处理负担。

下面简单小结三种交换技术的主要特点。

(1) 电路交换:在数据传输开始之前必须先设置一条完全的通路;在线路释放以前,该通路将被一对用户完全占用;对于猝发式的通信,电路交换效率不高。

(2) 报文交换:报文从源点传输到目的地采用存储转发的方式,在传输报文时,只占用一段通道;在交换节点中需要缓冲存储,报文需要排队,因此,报文交换不能满足实时通信的要求。

(3) 报文分组交换:交换方式和报文交换方式类似,但报文被分成分组传输,并规定了最大的分组长度;在数据报分组交换中,目的地需要重新组装报文;报文分组交换技术是数据网络中使用最广泛的一种交换技术。

6.3 IP 协议

IP 协议是 TCP/IP 体系中两个非常重要的协议之一,其定义了用以实现面向无连接服务的网络层分组格式,其中包括 IP 寻址方式。不同网络技术的主要区别在数据链路层和物理层,如不同的局域网技术和广域网技术。而 IP 协议则能够将不同的网络技术在 TCP/IP 的网络层统一在 IP 协议之下,以统一的 IP 分组传输提供对异构网络互联的支持。IP 协议使互联起来的许多计算机网络能够通信,因此,TCP/IP 体系中的网络层常常被称为网际层(Internet Layer),或 IP 层。

IP 协议是一个不可靠的、面向无连接的数据报传输协议。“面向无连接”表明 IP 协议不维护 IP 分组发送后的状态信息,并且每个数据分组的处理是相对独立的,“不可靠”意味着 IP 协议不能保证每个 IP 分组能够成功地到达目的节点,也不保证分组传输顺序的正确性。也就是说,IP 协议提供的是一种“尽力而为(Best-Effort)”的数据传输服务。相应地,在 IP 协议所定义的 IP 分组中,不提供任何恢复丢失或损坏数据分组这样的特殊功能,也不提供差错控制和确认机制。由于无须提供这些服务,因此网络可以更有效地运行。IP 协议的第一个功能是寻址。

6.3.1 IP 地址

1. IP 地址的结构、分类与表示

IP 地址以 32b 二进制位的形式存储于计算机中。32b 的 IP 地址结构由网络标识和主

机号两部分组成，如图 6.7 所示。其中，网络标识用于标识该主机所在的网络，而主机号则表示该主机在相应网络中的特定位置。

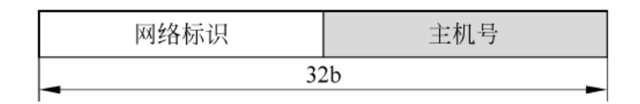


图 6.7　IP 地址的组成

目前，大多数 IP 编址方案仍采用 IPv4 编址方案，即使用 32b 二进制数组成，为了方便使用，将 IP 地址的 32b 二进制分成四段，每段 8b，中间用小数点隔开，然后将每 8b 二进制转换成十进制数，这种方法叫作点分十进制表示法。正是因为网络标识所给出的网络位置信息才使得路由器能够在通信子网中为 IP 分组选择一条合适的路径。

在现实世界中，网络根据需求的不同，有的网络可能含有较多的计算机，也有的网络可能包含较少的计算机，于是按照网络规模的大小，IP 协议将 IP 地址分成 A、B、C、D、E 5 类，其中 A 类、B 类、C 类作为普通的主机地址，D 类用于提供网络组播服务或作为网络测试之用，E 类保留给未来扩充使用，如图 6.8 所示。不同类别的 IP 地址的网络号和主机号的长度划分不同，它们能够识别的物理网络数不同，每个物理网络所能容纳的主机台数也不相同，如表 6.2 所示列出各类 IP 地址的特点。



图 6.8　IP 地址的组成

表 6.2　各类 IP 地址的特点

类别	类标识	第一字节	网络地址长度	主机地址长度	最大网络数	最大主机数	适用范围
A 类	0	1～126	1B	3B	126	16 777 214	大型网络
B 类	10	128～191	2B	2B	16 382	65 534	中型网络
C 类	110	192～223	3B	1B	2 097 150	254	小型网络
D 类	1110	224～239	—	—	—	—	多点播送
E 类	11110	240～247	—	—	—	—	保留地址

1) A 类地址

A 类地址结构如图 6.8 所示。A 类 IP 地址就由 1B 网络地址和 3B 主机地址组成，网

络地址的最高位必须是0。A类IP地址中网络标识长度为7b,最多可以提供126(2^7-2)个网络标识号,主机标识的长度为24b,A类网络地址数量非常多,可以用于主机数达1678($2^{24}-2$)万台的大型网络。

2) B类地址

B类地址结构如图6.8所示。B类IP地址就由2B网络地址和2B主机地址组成,网络地址的最高位必须是10。B类IP地址中网络标识长度为14b,最多可以提供16 384($2^{14}-2$)个网络标识号,主机标识的长度为16b,B类网络地址适用于中等规模的网络,每个网络所能容纳的计算机数为65 534($2^{16}-2$)台。

3) C类地址

C类地址结构如图6.8所示。C类IP地址就由3B网络地址和1B主机地址组成,网络地址的最高位必须是110。C类IP地址中网络标识长度为21b,最多可以提供200($2^{21}-2$)万个网络标识号,主机标识的长度为8b,C类网络地址数量较多,适用于小规模的局域网,每个网络最多只能包含254(2^8-2)台计算机。

4) D类地址

D类地址是多播地址,主要是留给Internet体系结构委员会IAB(Internet Architecture Board)使用,E类地址保留在今后使用。目前大量使用的IP地址仅有A类、B类、C类三类地址。

5) 特殊IP地址

除了上面介绍的IP地址外,还有几种特殊类型的IP地址,TCP/IP协议规定,凡IP地址中的第一个字节以“1110”开始的地址都叫作多点广播地址。因此,任何第一个字节大于223小于240的IP地址是多点广播地址;IP地址中的每一个字节都为0的地址(0.0.0.0)对应于当前主机,IP地址中的每一个字节都为1的IP地址(255.255.255.255)是当前子网的广播地址。

2. 特殊IP地址

关于最大网络数和每个网络可容纳的最大主机数目这两项上,都在相应的理论值后面减去了2,这是为什么呢?我们说,在IP地址空间中,凡是网络标识或主机号部分取值为全0和全1的地址都具有特殊的含义,被保留作为特殊之用。

1) 网络地址

当用户要表示一个网络时就要用到网络地址。在IP地址编码方案中,网络地址由一个有效的网络号和全0的主机号构成。如某主机的IP地址为168.36.12.55,这是一个B类地址,则此主机所在网络的地址为168.36.0.0。

2) 直接广播地址

当用户想向互联网的某个网络中所有主机发送数据报,叫作直接广播,具有这种特点的IP地址叫作直播广播地址。在IP地址编码方案中,直接广播地址由一个有效的网络号和全1的主机号构成。如当互联网中的一台主机使用168.36.255.255为目标地址发送数据报时,网络号为168.36.0.0的网络中所有主机都能收到该数据报。

3) 有限广播地址

当用户想向本网中每一台主机发送数据报,叫作有限广播。有限广播将广播限制在最小的范围内,当采用标准的IP地址编码,有限广播将发生在本网络之中,若采用子网编址,

有限广播将被限制在本子网中。有限广播地址为 255.255.255.255。

4）本网特定主机地址

当用户想与本网内部特定主机通信时，可通过将网络地址全部设为 0 进行简化（或不知道本网的网络地址）。如某主机发送数据报时，其目标 IP 地址为 0.0.136.32（B 类地址），则表示该数据报要送到本网主机号为 136.42 的主机上。

5）回送地址

A 类地址中，网络地址为 127 的地址用于网络软件测试或本机进程间通信。发送到这种地址的数据报不输出到线路上，立即返回。

除上述保留地址外，在 IPv4 的地址空间中，还保留了一部分被称为私有地址（Private Address）的地址资源，它们供企业、公司或组织机构内部组建 IP 网络时使用。私有地址包含了 A 类、B 类和 C 类地址空间中的 3 个小部分。它们分别是 A 类地址中的 10.0.0.0 ～ 10.255.255.255、B 类地址中的 172.16.0.0～172.31.255.255 和 C 类地址中的 192.168.0.0～192.168.255.255。根据规定，所有以私有地址为目标地址的 IP 数据报都不能被路由至外面的因特网上，否则就会违背 IP 地址在互联网环境中具有全局唯一性的约定。这些以私有地址作为逻辑标识的主机若要访问外面的因特网，必须采用网络地址翻译（Network Address Translation，NAT）或应用代理（Proxy）方式。

6.3.2 逻辑地址和物理地址

每一个物理网络中的网络设备都有其真实的物理地址。物理网络的技术和标准不同，其物理地址编码也不同。以太网物理地址用 48b 二进制数编码。因此可以用 12 个十六进制数表示一个物理地址。一般格式为 00-10-5a-63-aa-99。物理地址也叫 MAC 地址，它是数据链路层地址，即二层地址。

物理地址通常由网络设备的生产厂家直接烧入设备的网络接口卡的 EPROM 中，它存储的是传输数据时真正用来标识发出数据的源端设备和接收数据的目的端设备的地址。也就是说，在网络底层的物理传输过程中，是通过物理地址来标识网络设备的，这个物理地址一般是全球唯一的。物理地址只能够将数据传输到与发送数据的网络设备直接连接的接收设备上。对于跨越互联网的数据传输，物理地址不能提供逻辑的地址标识手段。

当数据需要跨越互联网时，使用逻辑地址标识位于远程目的地的网络设备的逻辑位置。通过使用逻辑地址，可以定位远程的节点。逻辑地址（如 IP 地址）则是第 3 层地址，所以有时又被称为网络地址，该地址是随着设备所处网络位置不同而变化的，即设备从一个网络被移到另一个网络时，其 IP 地址也会相应地发生改变。也就是说，IP 地址是一种结构化的地址，可以提供关于主机所处的网络位置信息。

总之，逻辑地址放在 IP 数据报的首部，而物理地址则放在 MAC 帧的首部。物理地址是数据链路层和物理层使用的地址，而逻辑地址是网络层和以上各层使用的地址。

6.3.3 IP 数据报

IP 分组由 IP 协议来定义。由于 IP 协议实现的是面向无连接的数据报服务，故 IP 分组

也称为 IP 数据报，相应地，IP 分组传输服务又被称为 IP 数据报服务。图 6.9 给出了 IP 分组的格式，可以看出，一个 IP 数据报由首部和数据两部分组成。首部的前一部分是固定长度，共 20B，是所有 IP 数据报必须具有的。在首部的固定部分的后面是一些可选字段，其长度是可变的。下面介绍首部各字段的意义。

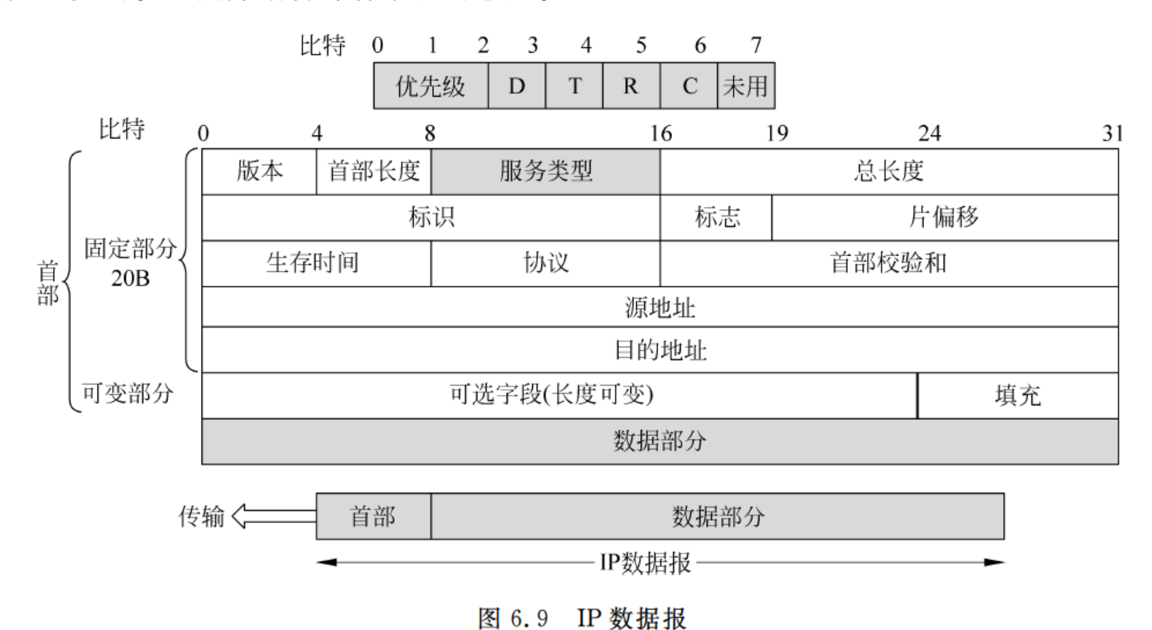


图 6.9 IP 数据报

(1) 版本：长度为 4b，表示数据报协议的版本。由于不同的协议版本所规定的数据报格式不同，所以必须提供一个字段用以说明数据报的版本信息，以免主机或路由设备上的 IP 软件在处理数据报时出现错误。目前的协议版本为 IPv4，下一代协议版本为 IPv6。

(2) 首部长度：长度为 4b，表示报头的长度。报头长度以 32b(相当于 4B)长度为一个单位。当报头中无可选项时，报头的基本长度为 5，相当于 20B；若一个报头有 IP 选项与填充字段，则报头长度要大于 5；报头长度的最大值为 15，即 60B。

(3) 服务类型：占 8b，主机要求通信子网提供的服务类型。包括一个 3b 长度的优先级，4 个标志位 D、T、R 和 C，D、T、R、C 分别表示延迟、吞吐量、可靠性和代价。另外，1b 未用。通常文件传输更注重可靠性，而数字声音或图像的传输更注重延迟。

(4) 总长度：占 16b，数据报的总长度，包括头部和数据，以字节为单位。数据报的最大长度为 $2^{16}-1$B，即 65 535B(即 64 KB)。

(5) 标识：占 16b，标识数据报。当数据报长度超出网络最大传输单元(MTU)时，必须进行分割，并且需要为分割段(Fragment)提供标识。所有属于同一数据报的分割段被赋予相同的标识值。在 IP 层下面的每一种数据链路层都有其自己的帧格式，其中包括帧格式中的数据字段的最大长度，称为最大传输单元(Maximum Transfer Unit，MTU)。当一个 IP 数据报封装成数据链路层的帧时，此数据报的总长度(即首部加上数据部分)一定不能超过下面的数据链路层的 MTU 值。例如，以太网的 MTU 为 1500B，FDDI 的 MTU 为 4352B，PPP 的 MTU 为 296B。在路由器接收到数据报并对其进行转发之前，必须从 MTU 的角度考虑转发接口所在的物理网络是否允许该数据报通过。

(6) 标志：占 3b，指出该数据报是否可分段。目前只有前两个比特有意义。

① 标志字段中的最低位记为 MF(More Fragment)。MF=1 即表示后面“还有分片”的

数据报。MF＝0 表示这已是若干数据报片中的最后一个。

② 标志字段中间的一位记为 DF(Don't Fragment)，表示不能分片。只有当 DF＝0 时才允许分片。

(7) 片偏移：占 13b，若有分段时，用以指出该分段在数据报中的相对位置，也就是说，相对于用户数据字段的起点，该片从何处开始。片偏移以 8B 为偏移单位，即每个分片的长度一定是 8B(64b)的整数倍。

例：该例子中数据报首部长度为 20B，数据区长度为 1600B，进入 MTU 为 1420B 的物理网络时进行第一次分片。第一次分片后，形成一个 1400B 的分片和一个 200B 的分片。第一片的片偏移为 0(0/8)，片未完标志为 1；第二片的片偏移为 175(1400/8)，片未完标志为 0，表示该片是数据报的最后一片。当第一个分片进入 MTU 为 820B 的物理网络时再次进行分片。第二次分片后，又形成了一个 800B 的分片和一个 600B 的分片。前者的片偏移为 0(0/8)，片未完标志为 1；后者的片偏移为 100(800/8)，片未完标志也为 1。分段过程如图 6.10 所示。

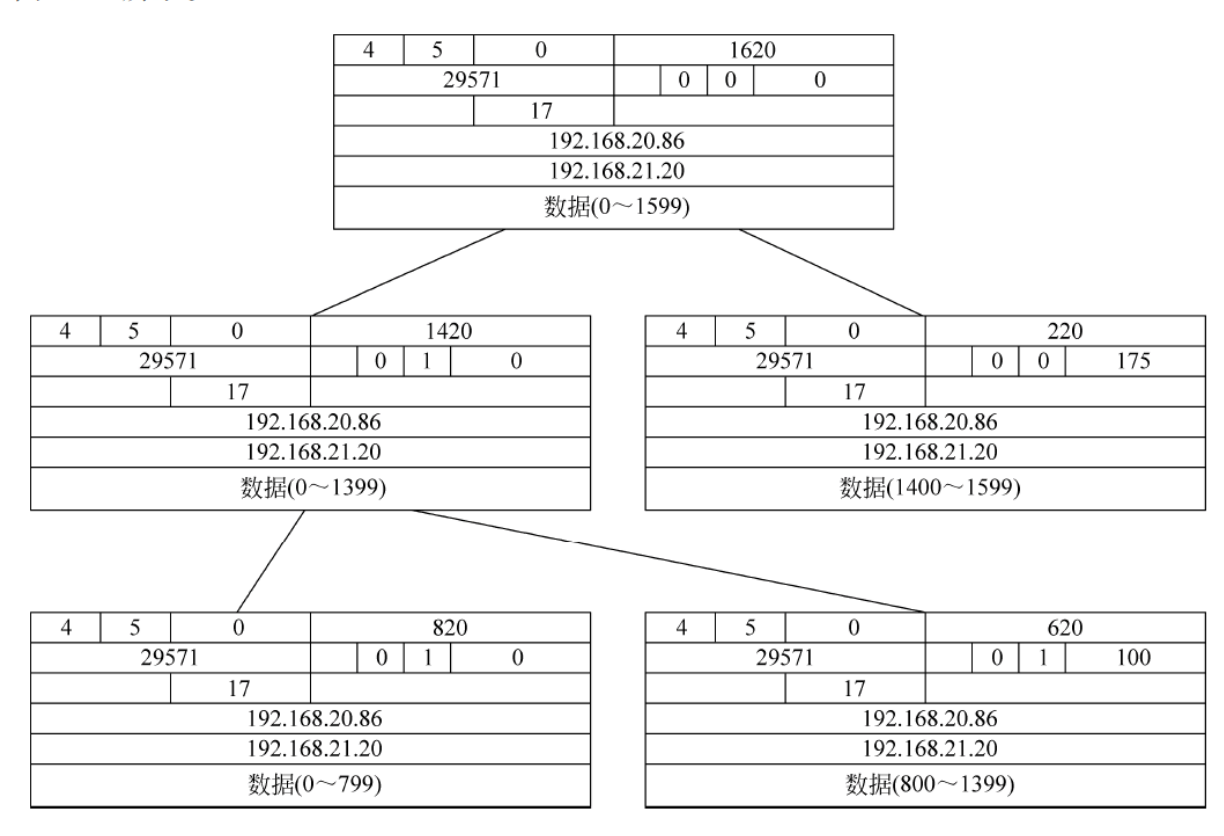


图 6.10 IP 数据报分片示例

(8) 生存时间或生命期：占 8b，记为 TTL(Time To Live)，即数据报在网络中的寿命，以秒来计数，建议值是 32s，最长为 $2^8-1=255$s。生存时间每经过一个路由节点都要递减，当生存时间减到零时，分组就要被丢弃。设定生存时间是为了防止数据报在网络中无限制地漫游。

(9) 协议：占 8b，指示传输层所采用的协议，如 TCP、UDP 或 ICMP 等。

(10) 首部校验和：占 16b，此字段只检验数据报的首部，不包括数据部分。采用累加求补再取其结果补码的校验方法。若正确到达时，校验和应为零。IP 协议对 IP 数据报首部

进行校验。原因如下。

① IP 首部属于 IP 层协议的内容,不可能由上层协议处理。

② IP 首部中的部分字段在点到点的传递过程中是不断变化的,只能在每个中间点重新形成校验数据,在相邻点之间完成校验。

IP 层不对数据进行校验。主要原因如下。

① 上层传输层是端到端的协议,进行端到端的校验比进行点到点的校验开销小得多,在通信线路较好的情况下尤其如此。

② 上层协议可以根据对于数据可靠性的要求,选择进行校验或不进行校验,甚至可以考虑采用不同的校验方法,这给系统带来了很大的灵活性。

(11) 可选字段:支持各种选项,提供扩展余地。根据选项的不同,该字段是可变长的,从 1B 到 40B。用来支持排错、测量以及安全等措施。

(12) IP 地址:占 32b,32b 的源地址与目的地址分别指出源主机和目的主机的网络地址。

IP 数据报中可选字段为 IP 数据报源站提供了两种显式路由信息的方法,它还为 IP 数据报提供了确定传输路由的方法。

(1) 不严格的源路由。不严格的源路由选项也称为不严格的源和记录路由选项,它为 IP 数据报提供了一种显式地提供路由信息的方法。路由器在把数据报转发到目的站时使用该信息,同时还用它记录路由。

(2) 严格的源路由。严格的源路由选项也称为 SSRR(Strict Source and Record Route,严格的源和记录路由)选项,除了中间路由器必须通过一个直接连接的网络把数据报发送到源路由中的下一个地址外,它使用与不严格的源路由相同的原则。它不能使用中间路由器。如果不实现这点,它就发出 ICMP 目的不可达的错误信息。

(3) 记录路由。这个选项提供了一种记录 IP 数据报通过路径的方法。它的功能类似于源路由选项。但是,这选项提供了一个空的路由数据字段,这个字段在数据报通过网络时被填入。源主机必须为这个路由信息提供足够的空间。如果数据字段在数据报到达目的主机之前被填充,则在不记录这个路径的情况下继续转发这个数据报。

(4) 时间戳。时间戳选项用于记录 IP 数据报经过各路由器时的当地时间,根据时间戳可以估算 IP 数据报从一台路由器到另一台路由器所花费的时间,从而帮助分析网络的吞吐率和负载情况。但由于大多数 IP 数据在 1s 的时间内就被转发及 IP 路由器不需要有同步的时钟,导致时间戳不精确。因此,它不能用于性能度量。

6.3.4　子网及子网划分

在 IP 地址规划时,常常会遇到这样的问题:一个企业或公司由于网络规模增加、网络冲突增加或吞吐性能下降等多种因素需要对内部网进行分段。而根据 IP 网络的特点,需要为不同的网段分配不同的网络号,于是当分段数量不断增加时,对 IP 地址资源的需求也随之增加。随着 Internet 规模的增大,32b 的 IP 地址空间已出现了严重的资源紧缺。

1. 子网概念

为了解决 IP 地址资源短缺的问题,同时也为了提高 IP 地址资源的利用率,引入了子网

划分技术。子网划分(Sub Networking)是指由网络管理员将一个给定的网络分为若干个更小的部分,这些更小的部分被称为子网(Subnet)。当网络中的主机总数未超出所给定的某类网络可容纳的最大主机数,但内部又要划分成若干个分段(Segment)进行管理时,就可以采用子网划分的方法。为了创建子网,网络管理员需要从原有 IP 地址的主机位中借出连续的若干高位作为子网络标识,如图 6.11 所示。也就是说,经过划分后的子网因为其主机数量减少,不需要原来那么多位作为主机标识,从而可以将这些多余主机位用作子网标识。



图 6.11 关于子网划分的示意图

子网编址将 IP 地址的主机号部分进一步划分成子网部分和主机部分,一个子网地址包括网络号、子网号和主机号 3 个部分。

2. 子网划分的方法

在子网划分时,首先要明确划分后所要得到的子网数量和每个子网中所要拥有的主机数,然后才能确定需要从源主机位借出的子网标识位数。原则上,根据全 0 和全 1IP 地址保留的规定,子网划分时至少要从主机位的高位中选择两位作为子网络位。A、B、C 类网络最多可借出的子网络位是不同的,A 类可达 22b、B 类为 14b,C 类则为 6b。显然,当借出的子网标识位数不同时,相应可以得到的子网数量及每个子网中所能容纳的主机数也是不同的。表 6.3 给出了子网位数、子网数量和有效子网数量之间的对应关系。所谓有效子网是指除去那些子网位为全 0 或全 1 的子网后所留下的可用子网。

表 6.3 子网位数与子网数量、有效子网数量的对应关系

子网位数	子网数量	有效子网数量
1	$2^1=2$	2－2＝0
2	$2^2=4$	4－2＝2
3	$2^3=8$	8－2＝6
4	$2^4=16$	16－2＝14
5	$2^5=32$	32－2＝30
6	$2^6=64$	64－2＝62
7	$2^7=128$	128－2＝126
8	$2^8=256$	256－2＝254
9	$2^9=512$	512－2＝510
…	…	…

下面以一个 C 类网络子网划分的例子来说明子网划分的具体方法。假设一个由路由器相连的网络,有 3 个相对独立的网段,并且每个网段的主机数不超过 30 台,现需要以子网划分的方法为其完成 IP 地址规划。由于该网络中所有网段合起来的主机数没有超出一个 C 类网络所能容纳的最大主机数,所以可以利用一个 C 类网络的子网划分来实现。假定为它们申请了一个 C 类网络 211.81.192.0,则在子网划分时需要从主机位中借出其中的高 3 位作为子网络位(思考为什么不能是 2 位),这样一共可得 8 个子网,每个子网的相关信息

参见表 6.4。其中,第 1 个子网因网络号与未进行子网划分前的原网络号 211.81.192.0 重复而不可用,第 8 个子网因为广播地址与未进行子网划分前的原广播地址 211.81.192.255 重复也不可用,这样可以选择 6 个可用子网中的任意 3 个为现有的 3 个网段进行 IP 地址分配,留下 3 个可用子网将作为未来网络扩充之用。

表 6.4　C 类地址 211.81.192.0 划分 8 个子网示例

子网的编号	借来的子网位的二进制数值	子网地址	子网广播地址	主机位可能的二进制数值(范围)(5 位)	子网/主机十进制数值的范围	是否可用
第 0 个子网	000	211.81.192.0	211.81.192.31	00000～11111	0～31	否
第 1 个子网	001	211.81.192.32	211.81.192.63	00000～11111	32～63	是
第 2 个子网	010	211.81.192.64	211.81.192.95	00000～11111	64～95	是
第 3 个子网	011	211.81.192.96	211.81.192.127	00000～11111	96～127	是
第 4 个子网	100	211.81.192.128	211.81.192.159	00000～11111	128～159	是
第 5 个子网	101	211.81.192.160	211.81.192.191	00000～11111	160～191	是
第 6 个子网	110	211.81.192.192	211.81.192.223	00000～11111	192～223	是
第 7 个子网	111	211.81.192.224	211.81.192.255	00000～11111	224～254	否

3. 子网划分的优越性

引入子网划分技术可以有效地提高 IP 地址的利用率,从而可节省宝贵的 IP 地址资源。在该例子中,假设没有子网划分技术,则至少需要申请 3 个 C 类地址,这样 IP 地址的使用率仅达 11.81%,而浪费率则高达 88.19%;采用子网划分技术后,尽管第 1 个和最后 1 个子网也是不可用的,并且在每个子网中又留出了一个网络号地址和广播地址,但 IP 地址的利用率却可以提高到 71%。表 6.5 给出了对一个 C 类网络进行不同位数的子网划分后所对应的 IP 地址利用率。

表 6.5　C 类网络子网划分后的 IP 地址利用率

所借位数(个)	创建的子网数(个)	每个子网中可拥有的主机数目(台)	可用的 IP 地址总数(个)	IP 地址利用率(%)
2	2	62	124	49
3	6	30	180	71
4	14	14	196	77
5	30	6	180	71
6	62	2	124	49

4. 子网划分举例

划分子网就是将一个大网络划分成几个较小的网络。A 类、B 类或 C 类 IP 地址都可以划分子网,划分子网是在 IP 地址编址的层次结构中增加了一个中间层次,使 IP 地址变成了三级层次结构。下面用 3 个例子,分别以 A 类、B 类和 C 类地址对子网规划与地址空间的划分方法进行说明。

1) A 类地址划分子网

一个 A 类地址是由 8b 网络号与 24b 主机号组成。如果一个组织得到一个完整 A 类

IP 地址，那么它就可以在一个单独的物理网络中为多达 16 777 214 台主机和路由器分配 IP 地址。但是，如果该组织希望有更多的物理网络，那么同样需要进行子网划分的工作。

【例 6.1】 一个大型公司管理者从网络管理中心 NIC 获得一个 A 类 IP 地址 121.0.0.0。该公司网至少需要由 1000 个子网组成(包括预留部分)。这些物理网络可以是 Ethernet、Token Ring、FDDI 或广域网。那么，网络设计者就有必要对这个 A 类 IP 地址进行子网划分。

设计满足这样一个用户需求的方案需要考虑以下几个问题。

(1) 该公司网需要 1000 个子网，加上子网号为全 0 和全 1 的两种特殊地址，那么子网的数量至少为 1002。

(2) 如果选择子网号位长为 9，那么子网总数最多可以达到 511，显然不能满足要求；如果选择子网号位长为 10，那么可以用来分配的子网数量最大可以为 1024，除去子网号为全 0 和全 1 这两个保留地址外，可以分配的子网数可以达到 1022 个，因此可以满足用户需要 1000 个子网号的要求。

在确定子网长度时，应该权衡子网数与每个子网中主机与路由器数这两个方面的因素，不能简单地追求子网的数量，一般是满足基本要求，并考虑留有一定的余量。

那么，对于一个 A 类地址，满足以上用户要求的子网划分方案是：网络号为 8b；子网号为 10b；主机号为 14b。A 类地址划分子网后的地址结构如图 6.12 所示。

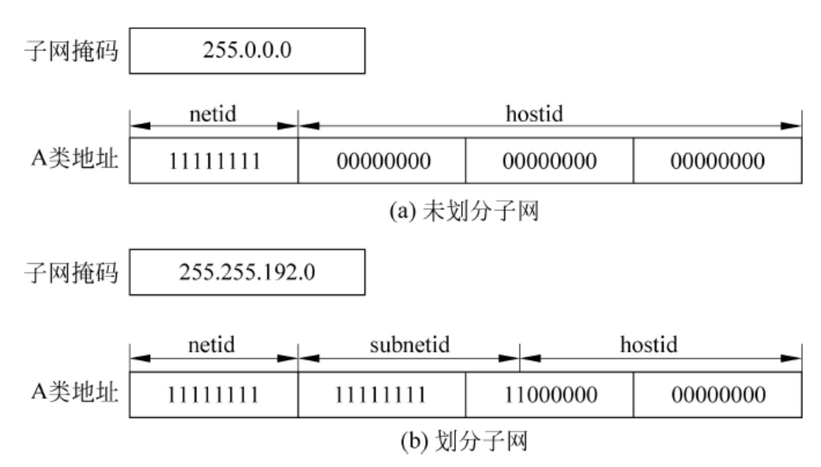


图 6.12 A 类地址划分子网后的地址结构

按照以上子网划分方法，A 类 IP 地址为 1022 个子网的各个子网的地址范围，如图 6.13 所示。除了特殊地址外，即子网号不允许是全 1 或全 0，其他地址均已在划分的子网中使用。

在以上子网划分的方案中，该企业网可以用的 IP 地址如下。

子网 1：121.0.64.1 ～ 121.0.127.254

子网 2：121.0.128.1 ～ 121.0.191.254

……

子网 1022：121.255.128.0 ～ 156.255.191.254

子网掩码：255.255.192.0

那么，在外部 Internet 用户看来，1000 个子网是作为一个整体出现的，它的网点地址为 121.0.0.0。

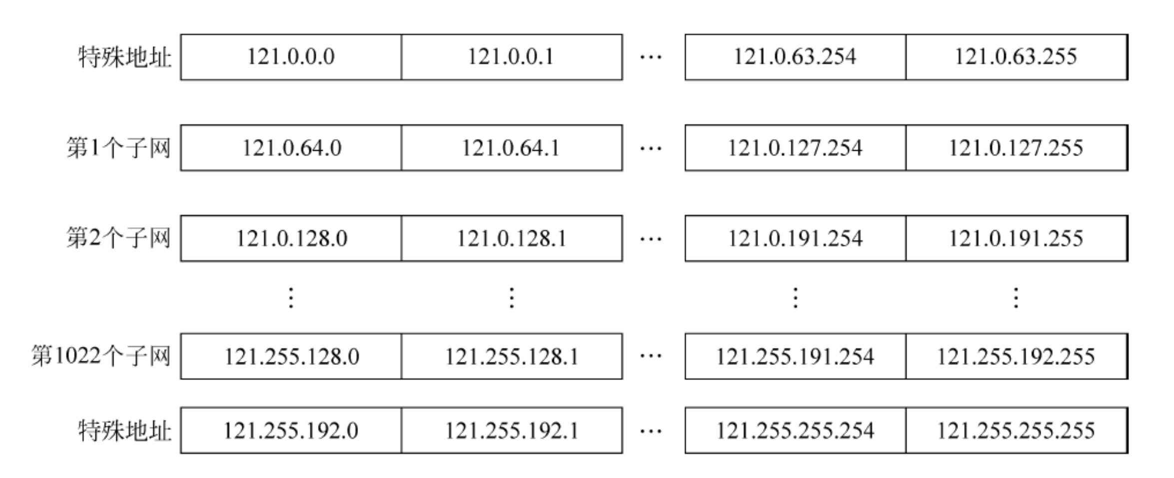


图 6.13 上例中的地址范围

2）B 类地址划分子网

一个 B 类地址是由 16b 网络号与 16b 主机号组成。如果一个组织得到一个 B 类 IP 地址，那么它可以在一个单独的网络中多达 65 523 个网络设备分配 IP 地址。但是，如果该组织希望有更多的物理网络，那么同样需要进行子网划分的工作。

【例 6.2】 一个校园网管理者从网络管理中心 NIC 获得一个 B 类 IP 地址：156.26.0.0。该校园网由近 210 个子网组成。这种结构显然是需要划分子网的。考虑到校园网的子网数量在 254 个之内，因此一个可行的方案是进行子网划分，子网号的长度为 8b。这样的子网掩码可以表示为 255.255.255.0。B 类地址划分子网后的地址结构如图 6.14 所示。

在以上子网划分的方案中，该校园网可以用的 IP 地址如下。

子网 1：156.26.1.1 ～ 156.26.1.254

子网 2：156.26.2.1 ～ 156.26.2.254

子网 3：156.26.3.1 ～ 156.26.3.254

……

子网 254：156.26.254.1 ～ 156.26.254.254

子网掩码：255.255.255.0

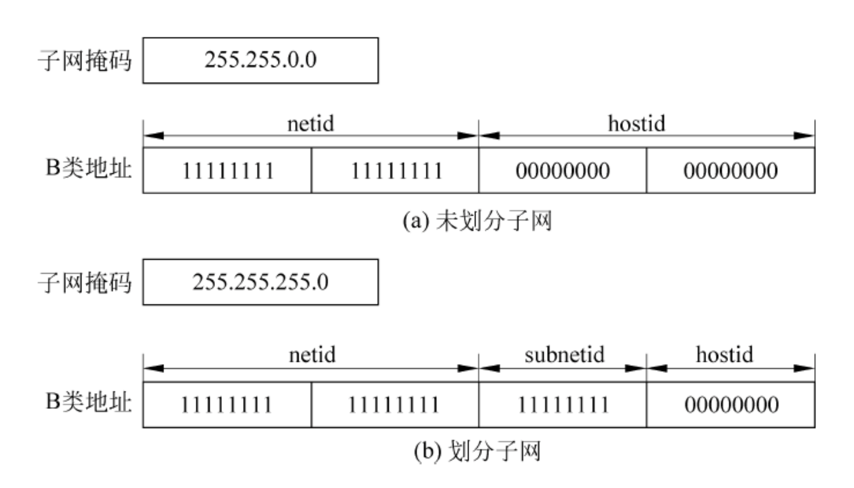


图 6.14 B 类地址划分子网后的地址结构

由于子网地址与主机号不能使用全 0 或全 1，因此校园网只能拥有 254 个子网，每个子网只能有 254 台主机。

【例 6.3】 如果例 6.2 中，校园网拥有的物理网络数超过 254 个，已有 900 个子网在使用。

显然，例 6.1 的方案也就不适用了。可以选择子网号位长为 10，子网个数总量可达到 1022 个，可以满足用户的要求。C 类地址划分子网后的地址结构如图 6.15 所示。这样，选择适用子网掩码是 255.255.255.192。那么子网划分后的 IP 地址的子网地址为 10b，主机号为 6b，其结构应该为：10b 的子网地址表示该校园网允许有 1022 个子网，6 位的主机号表示每个子网上可以有 62 台主机。

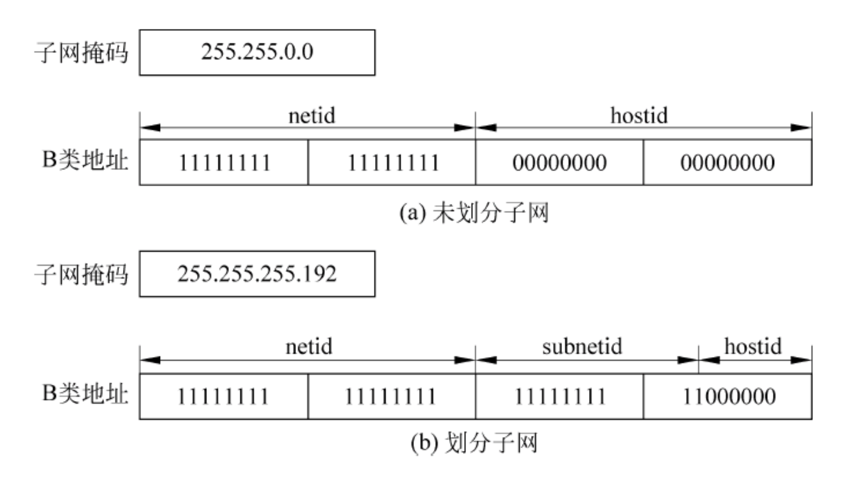


图 6.15 C 类地址划分子网后的地址结构

在以上子网划分的方案中，该校园网可以用的 IP 地址如下。

子网 1：156.26.0.65 ～ 156.26.0.126

子网 2：156.26.0.129 ～ 156.26.0.190

子网 3：156.26.0.193 ～ 156.26.0.254

子网 4：156.26.1.1 ～ 156.26.1.62

……

子网 1022：156.26.255.129 ～ 156.26.255.190

子网掩码：255.255.255.192

同样，由于子网号与主机号不允许是全 0 或全 1，因此总共可以使用的子网数为 1022 个。

3）C 类地址划分子网

一个 C 类地址是由 24b 网络号与 8b 主机号组成。如果一个单位得到一个 C 类 IP 地址，那么它可以在一个单独的网络中为 254 台主机与路由器分配 IP 地址。但是，如果该组织希望有更多的物理网络，那么同样需要进行子网划分的工作。

【例 6.4】 一个机关网管理者从网络管理中心 NIC 获得一个 C 类 IP 地址：212.26.220.0，该机关网是由 5 个子网组成。

该网络需要有 5 个子网，如果考虑到两个作为保留的特殊地址，那么需要子网号的总数为 7 个。显然，选择子网号位长为 3 即可满足用户要求。

C 类地址划分子网后的地址结构如图 6.16 所示。

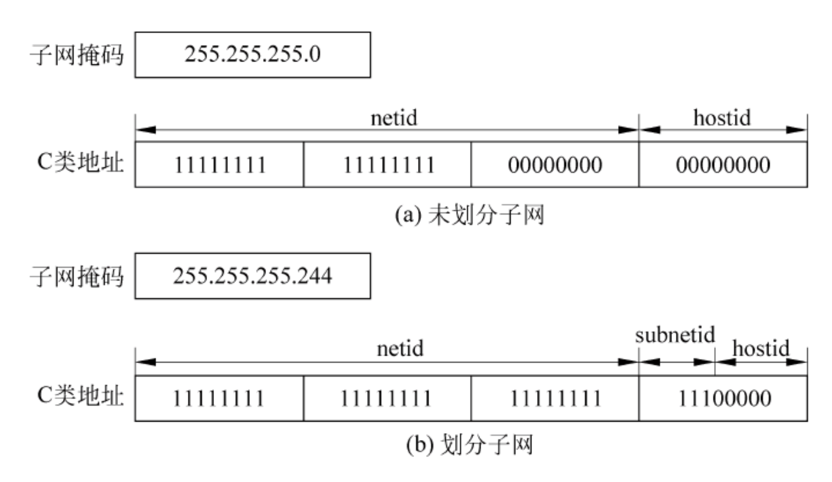


图 6.16 C 类地址划分子网后的地址结构

可以选择适用子网掩码 255.255.255.224,那么子网划分后的 IP 地址的子网号为 3b,主机号为 5b,其结构应该为 3b 的子网号表示该机关网允许有 6 个子网,5b 的主机号表示每个子网上可以有 30 台主机。

在以上子网划分的方案中,该机关网可以用的 IP 地址如下。

子网 1:212.26.220.33 ～ 212.26.220.62

子网 2:212.26.220.65 ～ 212.26.220.94

……

子网 6:212.26.220.193 ～ 212.26.220.222

子网掩码:255.255.255.224

6.3.5 子网掩码

前面讲过,网络标识对于网络通信非常重要。但引入子网划分技术后,带来的一个重要问题就是主机或路由设备如何区分一个给定的 IP 地址是否已被进行了子网划分,从而能正确地从中分离出有效的网络标识(包括子网络号的信息)。通常,将未引进子网划分前的 A 类、B 类、C 类地址称为有类别的(Classful)IP 地址,对于有类别的 IP 地址,显然可以通过 IP 地址中的标识位直接判定其所属的网络类别并进一步确定其网络标识。但引入子网划分技术后,这个方法显然是行不通了。例如,一个 IP 地址为 102.2.3.3,已经不能简单地将其视为一个 A 类地址而认为其网络标识为 102.0.0.0。因为若是进行了 8b 的子网划分,则其就相当于一个 B 类地址且网络标识成为 102.2.0.0;如果是进行了 16b 的子网划分,则又相当于一个 C 类地址且网络标识成为 102.2.3.0;若是其他位数的子网划分,则甚至不能将其归入任何一个传统的 IP 地址类中,可能既不是 A 类地址,也不是 B 类或 C 类地址。换言之,引入子网划分技术后,IP 地址类的概念已不复存在。对于一个给定的 IP 地址,其中用来表示网络标识和主机号的位数可以是变化的,这取决于子网划分的情况。将引入子网划分技术后的 IP 地址称为无类别的(Classless)IP 地址,并因此引入子网掩码的概念来描述 IP 地址中关于网络标识和主机号位数的组成情况。

子网掩码(Subnet Mask)通常与 IP 地址配对出现,其功能是告知主机或者路由设备,

一个给定 IP 地址的哪一部分代表网络号，哪一部分代表主机号。子网掩码采用与 IP 地址相同的位格式，由 32b 长度的二进制比特位构成，也被分为 4 个 8b 组并采用点十进制来表示。但在子网掩码中，所有与 IP 地址中的网络与子网位部分对应的二进制位取值为 1，而与 IP 地址中的主机位部分对应的位则取值为 0。

引入子网掩码后，不管是否进行过某种方式的子网划分，主机或路由器都可以通过将子网掩码与相应的 IP 地址进行求"与"操作，来提取出给定 IP 地址所属的网络号(包括子网号)信息。对主机来说，在发送一个 IP 数据报之前，它会通过将本机 IP 地址的子网掩码分别与源 IP 地址和目标 IP 地址进行求"与"操作，提取出相应的源网络号和目的网络号以判断源主机和目的主机是否在同一网络中。对路由设备而言，一旦从某一个接口接收到一个数据包，则会以该接口 IP 地址所对应的子网掩码与所收到的 IP 数据报里给出的目标 IP 地址进行求"与"操作，提取出目的网络号后再作为下一步路径选择的依据。

【例 6.5】 一个网络被分配了一个 C 类地址 202.113.27.0。如果该网络需要由 5 个子网组成，每个子网的计算机不超过 25 台，那么应该怎样规划和使用 IP 地址呢？其划分过程如下。

(1) 由于每个子网需要一个唯一的子网号来表示，即需要 5 个子网号。

(2) 因为每个子网的计算机不超过 25 台，考虑到使用路由器连接，因此需要至少 27 个主机号；可以分析出，选择子网掩码 255.255.255.224 可以满足要求，所对应的二进制地址是 1111111.11111111.11111111.11100000。

(3) 确定可用的网络地址：子网掩码确定后，可以确定使用的子网号位数。在本例中，子网号的位数为 3，因此可能的组合为 000、001、010、011、100、101、110 和 111。根据子网划分的规则，除去 000 和 111，剩余 001、010、011、100、101 和 110 6 个子网，因此所需 5 个子网的地址可分别选定为 202.113.27.32、202.113.27.64、202.113.27.96、202.113.27.128 和 202.113.27.160。

(4) 确定各个子网的主机地址范围，如表 6.6 所示。

表 6.6 各个子网对应的主机地址范围

子 网 地 址	主机地址范围
202.113.27.32	202.113.27.33～202.113.27.63
202.113.27.64	202.113.27.65～202.113.27.95
202.113.27.96	202.113.27.97～202.113.27.127
202.113.27.128	202.113.27.129～202.113.27.159
202.113.27.160	202.113.27.161～202.113.27.191

在很多情况下，需要根据两台主机的 IP 地址判断是否属于同一个物理子网。判断两台主机是不是在同一个子网中，其标准是看它们的子网地址是不是相同。在比较中需要将它们的地址用二进制形式表示。

例如，主机 1 与主机 2 的 IP 地址分别为 156.26.27.71、156.26.27.110，子网掩码为 255.255.255.192，判断它们是不是在同一个子网上。

解决方法是：首先用二进制方式写出它们的 IP 地址。

主机 1：10010010.00011010.00011011.01000111。

主机 2：10010010.00011010.00011011.01101110。

在一个子网中，所有的主机都具有相同的子网掩码。当知道网络中一台主机的 IP 地址与子网掩码，则将 IP 地址与子网掩码按位进行“与”(and)运算，其结果即为该主机所在子网的子网地址。可以将主机 1 的 IP 地址与子网掩码按位进行“与”(and)运算。

主机 1 的 IP 地址：　10010010.00011010.00011011.01000111

子网掩码：　　　　　11111111 .11111111. 11111111.11000000

“与”运算结果：　　10010010.00011010.00011011.01000000

由于该主机 1 的 IP 地址 156.26.27.71 是一个 B 类地址，因此它的前 16b 网络号，从“与”运算结果来看，它的子网号是 0001101101。

同样，也可以对主机 2 的 IP 地址与子网的二进制数，按位进行“与”(and)运算。

主机 1 的 IP 地址：　10010010.00011010.00011011.01000111

子网掩码：　　　　　11111111. 11111111. 11111111.11000000

“与”运算结果：　　10010010.00011010.00011011.01000000

从“与”运算结果来看，它的子网号也是 0001101101。这就说明：主机 1 与主机 2 的网络号与子网号都相同，因此它们属于同一个子网。

但是，也不是所有的 IP 地址在表面很相近的主机一定是属于同一个子网。例如，主机 3 与主机 4 的 IP 地址分别为 156.26.101.88、156.26.101.132，使用子网掩码也是 255.255.255.192。

首先用二进制方式写出它们的 IP 地址。

主机 3：10010010.00011010.01100101.01011000。

主机 4：10010010.00011010.01100101.10101110。

根据以上方法进行比较，发现主机 3 的子网地址为 0110010101，而主机 4 的子网地址为 0110010110。那么，尽管两者的网络号相同，但是 Sub 网络号不相同。因此，可以判断出两台主机不在同一个子网中。

在了解了子网划分的基本方法、主机 IP 地址与子网掩码，以及子网地址运算术方法之后，图 6.17 给出了一个子网划分的例子。从这个例子中也可以进一步理解子网划分的目的以及网络结构关系。

6.3.6　可变长子网掩码

当利用子网划分技术来进行 IP 地址规划时，常常会遇到各子网主机规模不一致的情况。例如，对一家企业或公司来说，可能在公司总部会有较多的主机，而分公司或部门的主机数会相对较少。为了尽可能地提高地址利用率，必须根据不同子网的主机规模来进行不同位数的子网划分，从而会在网络内出现不同长度的子网掩码长度并存的情况。我们将这种允许在同一网络范围内使用不同长度子网掩码的情况称为可变长子网掩码(Variable-Length Subnet Mask，VLSM)。

【例 6.6】　某个公司申请了一个整个 C 类 202.60.31.0 的 IP 地址空间。该公司有

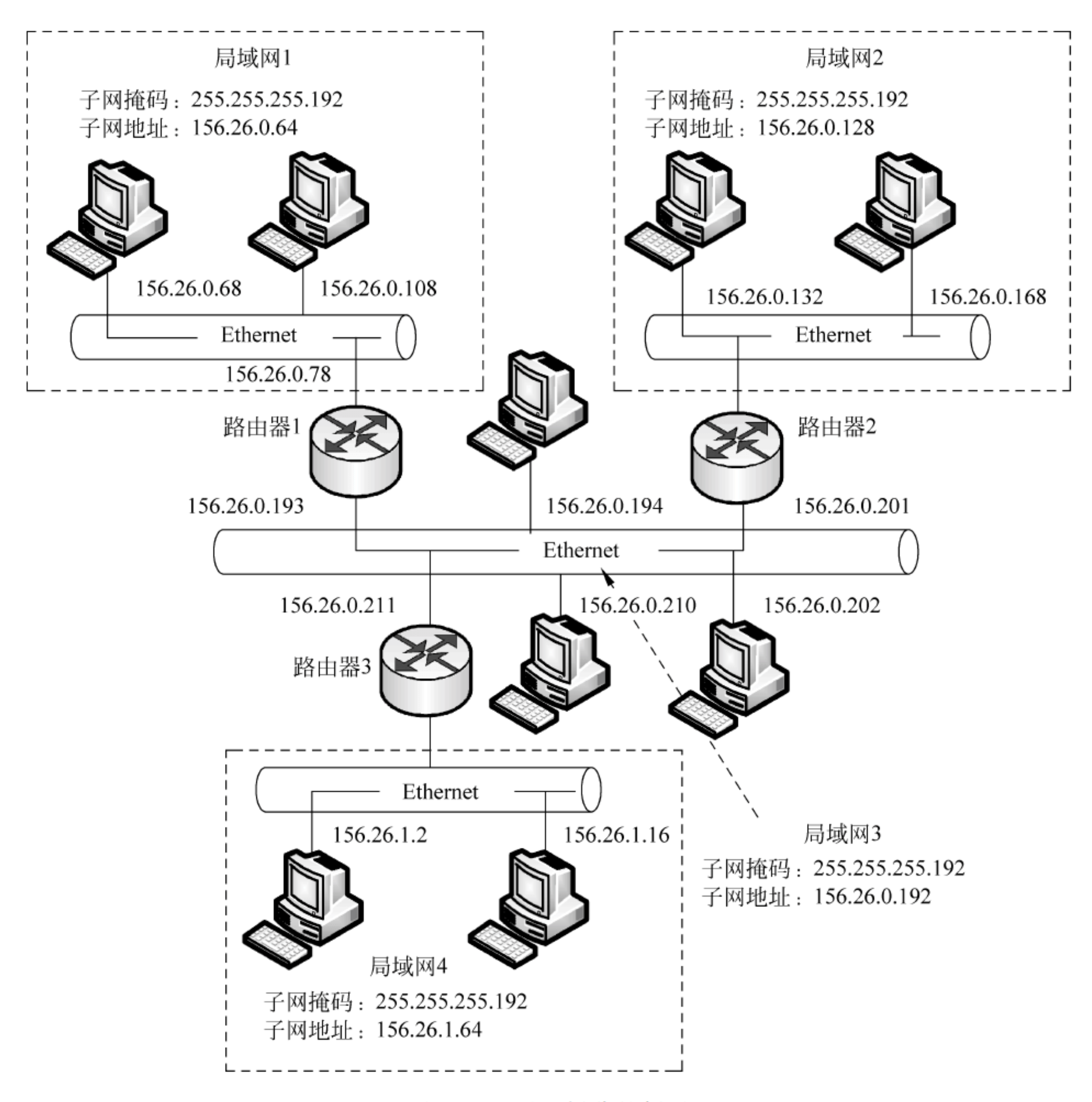


图 6.17　子网划分的例子

100 名员工在销售部门工作，50 名员工在财务部门工作，50 名员工在设计部门工作。要求为销售部门、财务部门与设计部门分别组建子网。

针对这种情况，可以通过可变长子网掩码技术，将一个 C 类 IP 地址分为 3 个部分，其中子网 1 的地址空间是子网 2 与子网3 的地址空间的两倍。那么，首先可以使用子网掩码 255.255.255.128，将一个 C 类 IP 地址划分为两半。在二进制计算中，运算过程如下。

主机的 IP 地址：　　11001010.00111100.00011111.00000000(202.60.31.0)

子网掩码：　　11111111.11111111.11111111.10000000(255.255.255.128)

“与”运算结果：　　11001010.00111100.00011111.10000000(202.60.31.0)

运算结果表明：可以将 202.60.31.11 ～ 202.60.31.126 作为子网 1 的 IP 地址，而将余下的部分进一步划分为两半。由于 202.60.31.127 第 4 个字节全是 1，被保留作为广播地址，不能使用，子网 1 与子网 2、子网 3 的地址空间交界点在 202.60.31.128，可以使用的子网掩码为 255.255.255.192。子网 2 与子网 3 的地址空间的计算过程如下。

主机的 IP 地址：	11001010.00111100.00011111.10000000(202.60.31.128)
子网掩码：	11111111. 11111111. 11111111.11000000(255.255.255.192)
“与”运算结果：	11001010.00111100.00011111.10000000(202.60.31.128)

现在可以将平分后的两个较小的地址空间分配给子网 2 与子网 3。对于子网 2 来说，第 1 个可用的地址是 202.60.31.129，最后一个可用的地址是 202.60.31.190。子网 2 的第 1 个可用的地址是 202.60.31.129 ～ 202.60.31.190。

因为下一个地址 202.60.31.191 中 191 是全 1 的地址，需要留作广播地址。接下来的一个地址是 202.60.31.192，它是子网 3 的第 1 个地址。那么，子网 3 的 IP 地址应该是 202.60.31.193 ～ 202.60.31.254。所以，采用可变长子网划分的 3 个子网的 IP 地址分别如下所示。

子网 1：202.60.31.1 ～ 202.60.31.126。

子网 2：202.60.31.129 ～ 202.60.31.190。

子网 3：202.60.31.191 ～ 202.60.31.254。

其中：子网 1 使用的子网掩码为 255.255.255.128，它们可以使用的 IP 地址数为 126 个；子网 2 与子网 3 的子网掩码为 255.255.255.192，它们可以使用的 IP 地址数分别为 61 个。该方案可以满足公司的要求。采用可变长子网掩码(VSLM)技术后，该公司网络的逻辑结构如图 6.18 所示。可变长子网划分的关键是找到合适的可变长子网掩码。

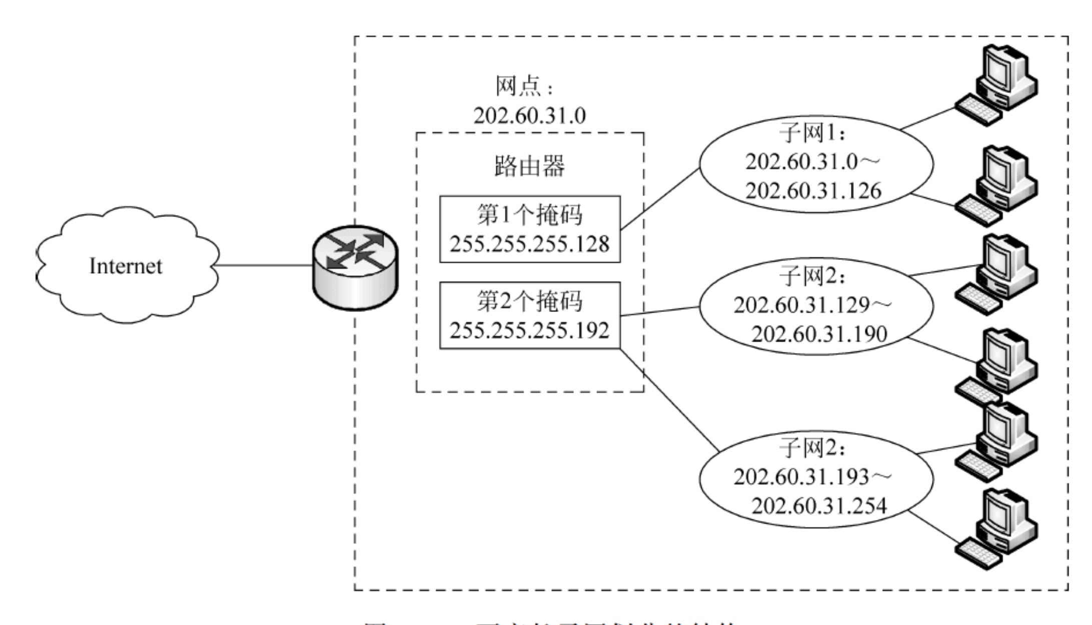


图 6.18　可变长子网划分的结构

6.3.7　地址解析协议

为使设备之间能够相互通信，源设备需要目的设备的 IP 地址和 MAC 地址。当一台设备试图与另一台已知 IP 地址的设备通信时，它必须确定对方的 MAC 地址。使用 TCP/IP 协议集中的地址解析协议(Address Resolution Protocol，ARP)可以自动获得 MAC 地址。ARP 允许主机根据 IP 地址查找 MAC 地址。每一台主机都设有一个 ARP 高速缓存，高速

缓存中有所在局域网上的各主机和路由器的 IP 地址到硬件地址的映射表。下面以图 6.19 所示的网络为例说明 ARP 的工作原理。

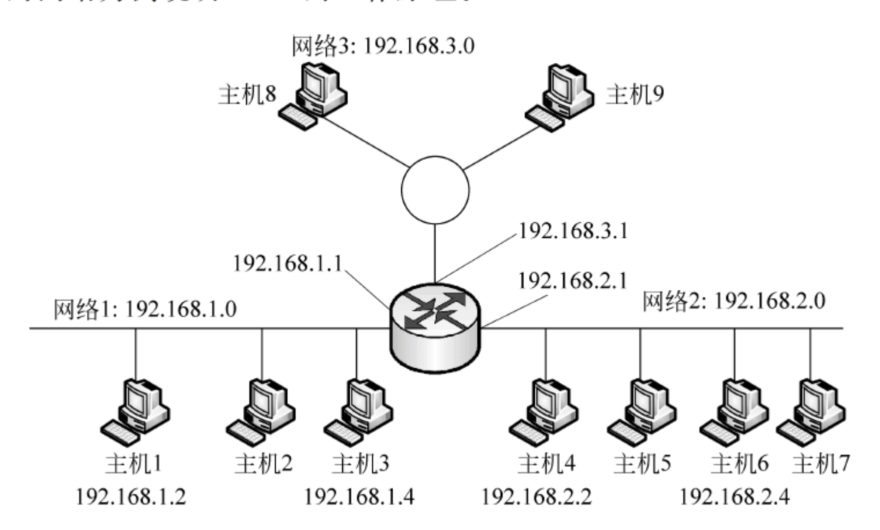


图 6.19　一个由路由器互联的网络

1. 子网内 ARP 解析

一台计算机能够解析另一台计算机地址的条件是这两台计算机都连在同一个物理网络中，如主机 1 向主机 3 发送数据报。主机 1 以主机 3 的 IP 地址为目的 IP 地址，以自己的 IP 地址为源 IP 地址封装了一个 IP 数据报；在数据报发送以前，主机 1 通过将子网掩码和源 IP 地址及目的 IP 地址进行求“与”操作判断源和目的在同一网络中；于是主机 1 转向查找本地的 ARP 缓存，以确定在缓存中是否有关于主机 3 的 IP 地址与 MAC 地址的映射信息；若在缓存中存在主机 3 的 MAC 地址信息，则主机 1 的网卡立即以主机 3 的 MAC 地址为目的 MAC 地址、以自己的 MAC 地址为源 MAC 地址进行帧的封装并启动帧的发送；主机 3 收到该帧后，确认是给自己的帧，进行帧的拆封并取出其中的 IP 分组交给网络层去处理。若在缓存中不存在关于主机 3 的 MAC 地址映射信息，则主机 1 以广播帧形式向同一网络中的所有节点发送一个 ARP 请求（ARP Request），在该广播帧中 48b 目的 MAC 地址以全 1 即 ffffffffffff 表示，并在数据部分发出关于“谁的 IP 地址是 192.168.1.4”的询问，这里 192.168.1.4 代表主机 3 的 IP 地址。网络 1 中的所有主机都会收到该广播帧，并且所有收到该广播帧的主机都会检查一下自己的 IP 地址，但只有主机 3 会以自己的 MAC 地址信息为内容给主机 1 发出一个 ARP 回应（ARP Reply）。主机 1 收到该回应后，首先将其中的 MAC 地址信息加入本地 ARP 缓存中，然后启动相应帧的封装和发送过程。

2. 子网间 ARP 解析

源主机和目的主机不在同一网络中，例如主机 1 向主机 4 发送数据报，假定主机 4 的 IP 地址为 192.168.2.2。这时若继续采用 ARP 广播方式请求主机 4 的 MAC 地址是不会成功的，因为第 2 层广播（在此为以太网帧的广播）是不可能被第 3 层设备路由器转发的。于是需要采用一种被称为代理 ARP（Proxy ARP）的方案，即所有目的主机不与源主机在同一网络中的数据报均会被发给源主机的默认网关，由默认网关来完成下一步的数据传输工作。注意，所谓默认网关是指与源主机位于同一网段中的某个路由器接口的 IP 地址，在此例中相当于路由器的以太网接口 F0/0 的 IP 地址，即 192.168.1.1。也就是说，在该例中，主机 1

以默认网关的 MAC 地址为目的 MAC 地址，而以主机 1 的 MAC 地址为源 MAC 地址，将发往主机 4 的分组封装成以太网帧后发送给默认网关，然后交由路由器来进一步完成后续的数据传输。实施代理 ARP 时需要在主机 1 上缓存关于默认网关的 MAC 地址映射信息，若不存在该信息，则同样可以采用前面所介绍的 ARP 广播方式得知，因为默认网关与主机 1 是位于同一网段中的。

6.3.8 反向地址解析协议

反向地址解析协议(RARP)把 MAC 地址绑定到 IP 地址上。这种绑定允许一些网络设备在把数据发送到网络之前对数据进行封装。一个网络设备或工作站可能知道自己的 MAC 地址，但是不知道自己的 IP 地址。设备发送 RARP 请求，网络中的一台 RARP 服务器出面来应答 RARP 请求，RARP 服务器有一个事先做好的从工作站硬件地址到 IP 地址的映射表，当收到 RARP 请求分组后，RARP 服务器就从这张映射表中查出该工作站的 IP 地址，然后写入 RARP 响应分组，发回给工作站。

6.4 IP 路由

6.4.1 路由选择基本原理

通信子网为网络中源节点到目的节点之间 IP 信息包的传递，提供多条传输路径，这样在网络节点上收到一个分组后，就要确定其下一个节点的传递路径，这个过程就是 IP 路由选择。它是网络层要实现的基本功能。

在数据报方式中，网络节点要为每一个分组做出路由选择；而在虚电路方式中，只需在建立连接的过程中确定路由。路由选择包括两个基本操作：最佳路径的判断、网间信息包的转发。其中，对于最佳路径的选择判断较为复杂，而在这里起传递作用的设备，最主要的就是路由器。

如图 6.20 所示，主机 178.16.0.1 到主机 192.168.0.1 就有多条路可走。所以，路径选择就成了路由器最重要的工作。

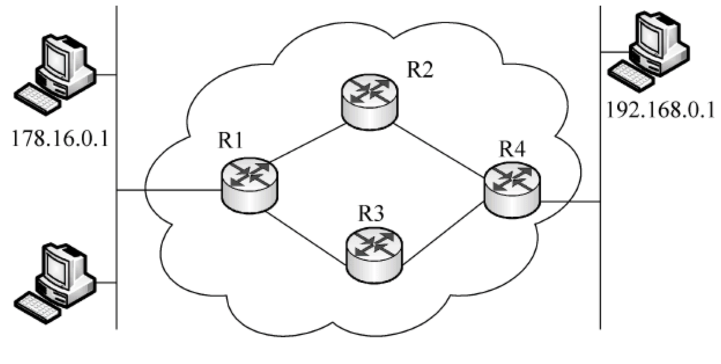


图 6.20 IP 路由选择

确定路由选择的策略被称为路由选择算法，它是路由选择的核心。首先，要考虑是选择最短路径还是最佳路径；其次，要考虑通信子网采用的是虚电路方式还是数据报方式；再次，还要考虑是集中式路由选择还是分布式路由选择；最后，是选择静态路由还是选择动态路由。路由器就是互联网中的中转站，网络中的数据包通过路由器转发到目的网络。在路由器的内部都有一张路由表，这张路由表中包含有该路由器掌握的目的网络地址以及通过此路由器到达这些网络的最佳路径，如某个接口或下一跳的地址，正是由于路由表的存在，路由器可以依据它进行转发。

路由器的某一个接口在接收到数据包以后，通过子网掩码求"与"运算从 IP 分组中提取出目标网络号，并将目标网络号与路由表进行比对看能否找到一种匹配，即确定是否存在一条到达目标网络的最佳路径信息。若不存在匹配，则将相应的 IP 分组丢弃。若存在匹配，又进一步地分成两种情况：第一种情况是路由器发现目的主机就在其直接相连的某个网络中，如图 6.21 中的主机 X，此时路由器就会去查找该目标 IP 地址所对应的 MAC 地址信息，并利用该地址信息将 IP 分组重新封装成目标网络所期望的帧发送到直接相连的目标网络中，这种形式的分组转发又被称为直接路由(Direct Routing)；第二种情况是路由器无法定位最后的目标网络，也就是说目的主机并不在路由器直接相连的任何一个网络中，但是路由器可以从路由表中找到一条与目标网络相匹配的最佳路径信息，如路由器转发接口的信息或下一跳路由器的 IP 地址等，如图 6.21 中的主机 Y，在这种情况下，路由器需要将 IP 分组重新进行封装成输出端口所期望的帧转发给下一跳路由器，由下一跳路由器继续后续的分组转发，这种形式的分组转发又被称为间接路由(Indirect Routing)。

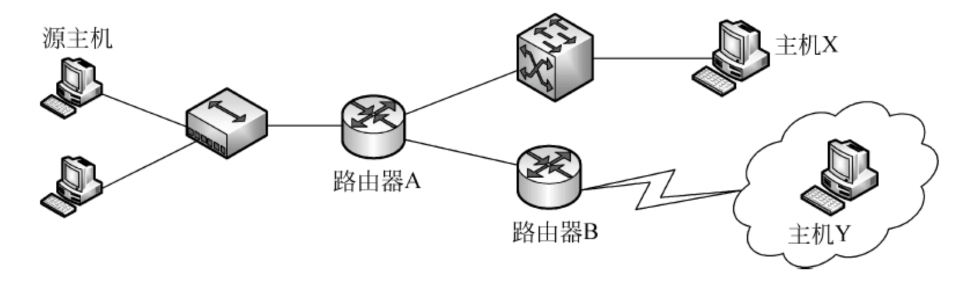


图 6.21 分组的直接路由和间接路由

当路由器从某个接口中收到一个数据包时，路由器查看数据包中的目的网络地址，如果发现数据包的目的地址不在接口所在的子网中，路由器查看自己的路由表，找到数据包的目的网络所对应的接口，并从相应的接口转发出去。上述描述是最基本的路由原理。

对路由器而言，上述查找路由表以获得最佳路径信息的过程被称为路由器的"路由"功能，而将从接收端口进来的数据在输出端口重新转发出去的功能称为路由器的"交换"功能。"路由"与"交换"是路由器的两大基本功能。

6.4.2 标准路由选择

1. 标准路由选择的基本构成

在 IP 互联网中，需要进行路由选择的设备一般采用表驱动的路由选择算法。每台需要路由选择的设备需要保存一份 IP 路由表，该表保存着可能到达的目的地址以及如何到达目

的网络的路径信息。在进行路由选择时，就会查询路由表，决定数据投递的路径。

路由表的基本结构一般由目的网络的 IP 地址，以及到下一个网络所必经的路由器 IP 地址组成。由于互联网中拥有大量主机，因而如果用具体的 IP 地址（网络号和主机号结合）来描述这些作为目的地的主机信息几乎不可能，所以我们往往借助于同一网络中所有主机共享的同一网络号，来作为目的地址的标识。总的来说有以下优点。

（1）可以大大减小路由表的规模。网络数比主机数要少得多，而信息到达信宿网络，也就到达了信宿主机，数据传递到信宿的相邻路由器后，相邻路由器再通过直接传递将数据传给信宿主机。所以可以用网络地址来取代网络中各主机的地址。

（2）与网络的抽象结构相对应。网络的抽象结构中只有网络，没有主机。

（3）增强了路由表对网络变化的适应性。由于体现了信息隐藏的原则，主机的增加和删去不会对路由表产生任何影响。

（4）减轻了路由表维护以及路由选择的开销，同时也简化了路由设备的设计和实现。

一个标准的 IP 路由表通常包含许多(N,R)对序偶结构，其中，N 代表目的网络的 IP 地址信息；R 代表要到网络 N 所经过的下一站路由器的 IP 地址信息。可以看出，就网络中的某一台路由器来说，它并不清楚到达目的地的完整路径信息，只是知道如何选择到达其下一站路由器的投递路径。在路由表中只采用下一跳地址而不用完整路径的好处如下。

（1）减小了路由表的规模。

（2）去掉了路由表中关于相同路径的冗余信息。

（3）使路由表变得简单，便于维护。

图 6.22 中，列举出了 3 台路由器各自的路由表。以路由器 B 为例，它与网络 10.0.2.0 和网络 10.0.3.0 直接相连，路由器 B 如果收到目的 IP 地址的网络号为 10.0.2.0 或 10.0.3.0，那么路由器 B 将该报文直接投递给目的主机。如果收到的目的地网络号为 10.0.1.0，那么路由器 R 就需要将该报文传送给与其相连的另外一台路由器 A，由路由器 A 再次转发该报文，直至目的地。以此如果接收到报文的目的地网络号为 10.0.4.0，则路由器 B 就需要再将该报文传送给路由器 C。

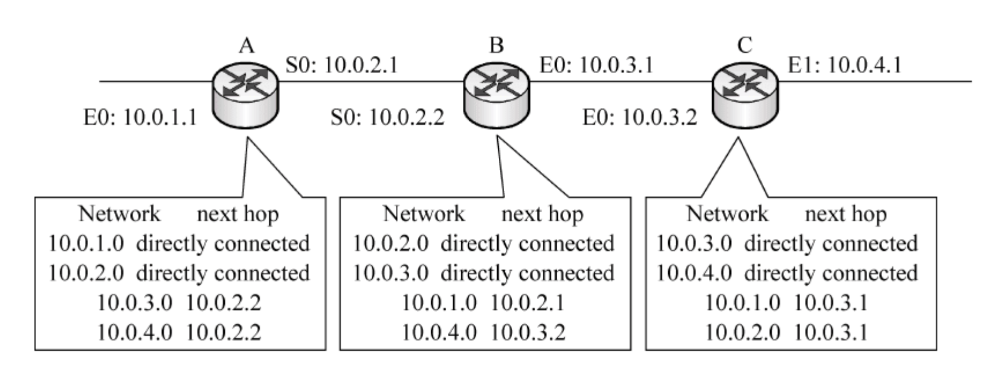


图 6.22 标准路由选择

2. 在主机上的路由表信息

在运行 Windows 系列操作系统的计算机上输入 route print 命令，即可以得到结构如下的本机路由表信息，如图 6.23 所示。

主机的 IP 路由表包含以下所列信息。

```
C:\>route print
===========================================================================
Interface List
0x1 ........................... MS TCP Loopback interface
0x2 ...00 50 8d 74 3c eb ...... IC Plus IP100 10/100 Fast Ethernet Adapter - Kas
persky Anti-Virus NDIS Miniport
===========================================================================
===========================================================================
Active Routes:
Network Destination        Netmask          Gateway       Interface  Metric
          0.0.0.0          0.0.0.0    10.112.10.254    10.112.10.99     20
      10.112.10.0    255.255.255.0     10.112.10.99    10.112.10.99     20
     10.112.10.99  255.255.255.255        127.0.0.1       127.0.0.1     20
   10.255.255.255  255.255.255.255     10.112.10.99    10.112.10.99     20
        127.0.0.0        255.0.0.0        127.0.0.1       127.0.0.1      1
        224.0.0.0        240.0.0.0     10.112.10.99    10.112.10.99     20
  255.255.255.255  255.255.255.255     10.112.10.99    10.112.10.99      1
Default Gateway:     10.112.10.254
===========================================================================
```

图 6.23 本机路由表信息

(1) 目标网络(Network Destination)。主机路由表里的目标可以是目的主机(10.112.10.99)、子网地址(10.112.10.0)、网络地址或默认路由(0.0.0.0)。路由表中的 127.0.0.0 代表本地环路网；224.0.0.0 代表组播网络。255.255.255.255 代表本地广播网。

(2) 网络掩码(Netmask)。网络掩码与目标位置结合使用以决定使用路由的时间。如主机路由的掩码为 255.255.255.255,默认路由的掩码为 0.0.0.0,而子网或网络路由的掩码在这两个极限值之间。掩码 255.255.255.255 表明只有精确匹配的目标位置使用此路由。掩码 0.0.0.0 表示任何目标位置都可以使用此路由。

(3) 网关(Geteway)。网关是数据包需要发送到下一台路由器的 IP 地址。本例中网关有 3 个：127.0.0.1 代表发往本地环回地址；10.112.10.254 是本机的默认网关地址,凡是不能在路由表中查找的目标,就发往本地址；10.112.10.99 是本机的 IP 地址,代表发往本机网卡。

(4) 接口(Interface)。接口表明用于接通下一台路由器的本地网卡的 IP 地址或本地环回地址。

(5) 度量参数(Metric)。本机的度量参数使用的是“跳数(跃点数)”,表明到达目标位置所通过的路由器数目。如果有多个到相同目标位置的跳数,则选择跳数最低的路由为最佳路径。

6.4.3 子网路由选择

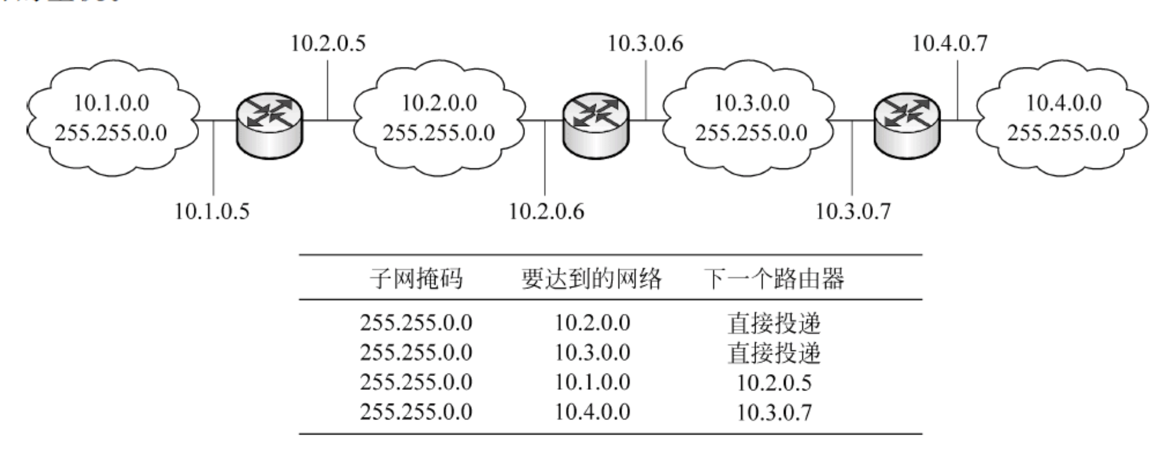
现实中很多网络不是采用标准路由的选路,而是采用子网路由选择路径。主要是因为很多网络是采用子网编制的结构,中间必须涉及子网掩码的区分,因此,必须更改标准路由选择方法,以满足子网选路的需求。

首先必须在标准路由表的对序偶(N,R)中加入区分子网的重要信息字段“子网掩码”,则子网的 IP 路由表表项最终可为(M,N,R)三元组结构。其中,M 代表子网掩码；N 代表目的网络地址；R 代表到达网络 N 所经过的下一站路由器的 IP 地址信息。

在子网路由选择时,首先将 IP 数据包中的目的地址 F 取出来并转换为二进制形式,接

着与路由表中的"子网掩码"进行逐位"与"运算，得到结果 F1 再与路由表中"目的网络地址"逐个进行比较，如果相同，说明路由选择成功，IP 数据报沿着对应的"下一站路由器 IP 地址"传送出去。

如图 6.24 所示，如果路由器 R 收到一个目的地址为 10.4.0.16 的数据报，那么它首先将该地址与路由表的第一项子网掩码 255.255.0.0 进行"与"运算，得到 F1=(10.4.0.0)，发现与第一项中 10.2.0.0 不符，说明路由选择不成功。这样它会按此方法逐项计算比较，直到最后一项。得到相与的结果 F1 与此项中 10.4.0.0 相一致，说明选路成功，这样路由器将此数据报转发给下一站路由器 S(10.3.0.7)，然后 S 会按照相同方法继续转发数据报直至目的主机。



子网掩码	要达到的网络	下一个路由器
255.255.0.0	10.2.0.0	直接投递
255.255.0.0	10.3.0.0	直接投递
255.255.0.0	10.1.0.0	10.2.0.5
255.255.0.0	10.4.0.0	10.3.0.7

图 6.24 子网路由

6.4.4 静态路由和动态路由

在路由器中维持一个能正确反映网络拓扑与状态信息的路由表对于路由器完成路由功能是至关重要的。通常有两种方式可用于路由表信息的生成和维护，分别是静态路由和动态路由。

所谓静态路由是指网络管理员根据其所掌握的网络连通信息以手动配置方式创建的路由表表项。这种方式要求网络管理员对网络的拓扑结构和网络状态有着非常清晰的了解，而且当网络连通状态发生变化时，静态路由的更新也要通过手动配置方式完成。静态路由通常作为网络测试、网络安全或带宽管理的有效措施。每台路由器中的静态路由表是一个本地文件，该文件包含所有去往已知网络的路由。静态路由要求手动配置固定的路由表。当网络结构发生变化时，网络管理员要及时地调整路由表。

当网络互联规模增大或网络中的变化因素增加时，依靠手动配置方式生成和维护一张路由表会变得不可想象，同时静态路由也很难及时适应网络状态的变化。此时希望有一种能自动适应网络状态变化而对路由表信息进行动态更新和维护的路由生成方式，这就是动态路由。动态路由是指路由协议通过自主学习而获得的路由信息，通过在路由器上运行路由协议并进行相应的路由协议配置即可保证路由器自动生成并维护正确的路由信息。使用路由协议动态构建的路由表不仅能更好地适应网络状态的变化，如网络拓扑和网络流量的变化，同时也减少了人工生成与维护路由表的工作量。但为此付出的代价则是用于运行路

由协议的路由器之间交换和处理路由更新信息而带来的资源耗费，包括网络带宽和路由器资源的占用。

6.4.5 路由协议

在网络层用于动态生成路由表信息的协议被称为路由协议，路由协议使得网络中的路由设备能够相互交换网络状态信息，从而在内部生成关于网络连通性的映像(Map)并由此计算出到达不同目的网络的最佳路径或确定相应的转发端口。

路由协议有时又被称为主动路由(Routing)协议，这是与规定网络层分组格式的网络层协议(如IP协议)相对而言的。IP协议的作用是规定了包括逻辑寻址信息在内的IP数据报格式，其使网络上的主机有了一个唯一的逻辑标识，并为从源到目的的数据转发提供了所必需的目标网络地址信息。但IP数据报只能告诉路由设备数据报要往何处去(What destination or Where to go)，还不能解决如何去的问题(How to reach)，而路由协议则恰恰提供了关于如何到达既定目的的路径信息。也就是说，路由协议为IP数据报到达目的网络提供了路径选择服务，而IP协议则提供了关于目的网络的逻辑标识并且是路由协议进行路径选择服务的对象，所以在此意义上又将IP协议这类规定网络层分组格式的网络层协议称为被动路由(Routed)协议。

路由协议的核心是路由选择算法。它为路由表中最佳路径的产生提供了算法依据。有多种不同的路由选择算法，通常，评价一个算法的优劣要考虑下面一些因素。

(1) 正确性。沿着路由表所给出的路径，分组一定能够正确无疑地到达目标网络或目的主机。

(2) 简单性。在保证正确性的前提下，路由选择算法要尽可能地简单，以减少最佳路径计算的复杂度和相应的资源消耗，包括路由器的CPU资源和网络带宽资源等。

(3) 健壮性。具备适应网络拓扑和通信量变化的足够能力。当网络中出现路由器或通信线路故障时，算法能及时改变路由以避免数据包通过这些故障路径；当网络中的通信流量发生变化时，如某些路径发生拥塞时，算法能够自动调整路由，以均衡网络链路中的负载。

(4) 稳定性。当网络拓扑发生变化时，路由选择算法能够很快地收敛。即网络中的路由器能够很快地捕捉到网络拓扑的变化，并在最快时间内对到达目标网络的最佳路径有新的一致认识或选择。

(5) 最优性。相对于用户所关心的那些开销因素，算法所提供的最佳路径确实是一条开销最小的分组转发路径。但是，由于不同的路由选择算法通常会采用不同的评价因子及权重来进行最佳路径的计算，因此在不同的路由选择算法之间，事实上并不存在关于最优的严格可比性。

路由选择算法在计算最佳路径时所考虑的因素被称为评价因子(Metric)。常见的评价因子包括带宽、可靠性、延时、负载、跳数和费用等。带宽是指通信链路的数据传输速率，通常情况下，一条高带宽的通信链路要优于一条相对低带宽的通信链路；可靠性是指数据传输过程中的质量，通常用误码率来表示；延时(Delay)是指一个分组从源主机到达目的主机所需的时间，延时与分组所经过的网络链路的带宽、负载及所经过的路由器性能都有关系；

跳数(Hop Count)是指从源主机到目的主机所需经过的路由器数目；费用(Cost)是指为了传输分组所付出的链路费用，通常这部分费用是由租用链路引起的。

通常，按路由选择算法的不同，路由协议被分为距离矢量路由协议、链路状态路由协议和混合型路由协议3大类。

表6.7给出了距离矢量路由协议与链路状态路由协议的比较。距离矢量路由协议的典型例子包括路由消息协议(Routing Information Protocol，RIP)和内部网关路由协议(Interior Gateway Routing Protocol，IGRP)等，链路状态路由协议的典型例子则是开放最短路径优先协议(Open Shortest Path First，OSPF)。混合型路由协议是综合了距离矢量路由协议和链路状态路由协议的优点而设计出来的路由协议，如IS-IS(Intermediate System-Intermediate System)和增强型内部网关路由协议(Enhanced Interior Gateway Routing Protocol，EIGRP)就属于此类路由协议。

表6.7　距离矢量路由协议与链路状态路由协议的比较

距离矢量路由协议	链路状态路由协议
从网络邻居的角度观察网络拓扑结构	得到整个网络的拓扑结构图
路由器转换时增加距离矢量	计算出通往其他路由器的最短路径
频繁、周期地更新；慢速收敛	由事件触发来更新；快速收敛
把整张路由表发送到相邻路由器	只把链路状态路由选择的更新传输到其他路由器上

按照作用范围和目标的不同，路由协议还可被分为内部网关协议和外部网关协议。内部网关协议(Interior Gateway Protocol，IGP)是指作用于自治系统以内的路由协议；而外部网关协议(Exterior Gateway Protocol，EGP)则是指作用于不同自治系统之间的路由协议。所谓自治系统(Autonomous System，AS)，是指网络中那些由相同机构操纵或管理，对外表现出相同的路由视图的路由器所组成的系统。自治系统由一个16b长度的自治系统号进行标识，其由NIC指定并具有唯一性。内部网关协议和外部网关协议的主要区别在于其工作目标的不同，前者关注于如何在一个自治系统内提供从源到目的的最佳路径，而外部网关协议则更多关注于能够为不同自治系统之间通信提供多种路由策略。RIP、IGRP、OSPF、EIGRP等都属于内部网关协议，在Internet上广为使用的边界网关协议(Border Gateway Protocol，BGP)则是外部网关协议的典型例子。

6.4.6　无分类编址

划分子网在一定程度上缓解了因特网在发展中遇到的困难。在一些比较大的企业的网络应用过程中，可以通过VLSM技术为它们部门或者分公司等划分子网。这样做有以下优点，一是可以提高IPv4地址的利用率；二是减小广播域的大小，降低产生广播风暴的可能；三是可以提高网络的安全性。但与此同时，这样进行子网划分也会带来一些负面影响，如为了使各子网互通，在各路由器的路由表中就可能添加许多子网级的数据项，使得整个网络的路由表变得更为复杂，影响整个网络的路由性能。1994年，在VLSM的基础上又进一步研究出无分类编址方法，它的正式名字是无分类域间路由选择CIDR(Classless Inter-Domain

Routing)。

1. 表示方法

IP 地址从三级编址(使用子网掩码)又回到了两级编址:

无分类的两级编址的记法是:IP 地址::= {<网络前缀>,<主机号>}

CIDR 还使用“斜线记法”(Slash Notation),它又称为 CIDR 记法,即在 IP 地址后面加上一个斜线“/”,然后写上网络前缀所占的比特数(这个数值对应于三级编址的子网掩码中比特 1 的个数)。

2. CIDR 的特点

(1) CIDR 消除了传统的 A 类、B 类和 C 类地址以及划分子网的概念,因而可以更加有效地分配 IPv4 的地址空间。

(2) IP 地址从三级编址(使用子网掩码)回到两级编址。

(3) CIDR 使用各种长度的“网络前缀”(Network-Prefix)来代替分类地址中的网络号和子网号。

我们只要知道 CIDR 地址块中的任何一个地址,就可以知道这个地址块的起始地址(即最小地址)和最大地址,以及地址块中的地址数。如图 6.25 所示,128.14.32.0/20 表示的地址块共有 2^{12} 个地址(因为斜线后面的 20 是网络前缀的比特数,所以主机号的比特数是 12),则这个地址块的起始地址是 128.14.32.0。128.14.32.0/20 地址块的最小地址是:128.14.32.0;128.14.32.0/20 地址块的最大地址是:128.14.32.255。当然,全 0 和全 1 的主机号地址一般不使用,通常只使用两个特殊地址之间的地址。在不需要指出地址块的起始地址时,也可将这样的地址块简称为“/20 地址块”。

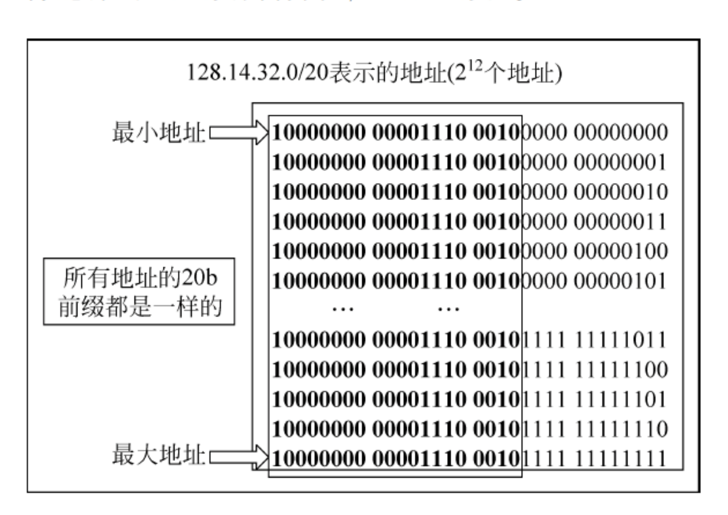


图 6.25 CIDR 地址块示例

为了更方便地进行路由选择,CIDR 使用 32b 的地址掩码(Address Mask)。地址掩码由一串 1 和一串 0 组成,而 1 的个数就是网络前缀的长度。虽然 CIDR 不使用子网了,但由于目前仍有一些网络还使用子网划分和子网掩码,因此 CIDR 使用的地址掩码也可继续称为子网掩码。例如,/20 地址块的地址掩码是:11111111 11111111 11110000 00000 000 (20 个连续的 1)。斜线记法中,斜线后面的数字就是地址掩码中 1 的个数。“CIDR 不使用子网”是指 CIDR 并没有在 32b 地址中指明若干位作为子网字段。但分配到一个 CIDR 地

址块的单位，仍然可以在本单位内根据需要划分出一些子网。这些子网也都只有一个网络前缀和一台主机号字段，但子网的网络前缀比整个单位的网络前缀要长一些。例如，某单位分配到地址块/20，就可以再继续划分为8个子网(即需要从主机号中借用3b来划分子网)。这时每一个子网的网络前缀就变成23b(原来的20b加上从主机号借来的3b)，比该单位的网络前缀多了3b。

斜线记法还有一个好处就是它除了表示一个IP地址外，可以很快地得到IP地址所在地址块的最小地址和最大地址。例如，地址192.199.170.82/27不仅表示IP地址是192.199.170.82，而且还表示这个地址块的网络前缀有27b(剩下的5b是主机号)，因此这个地址块包含32个IP地址($2^5=32$)。通常简单地计算还可得出，这个地址块的最小地址是192.199.170.64，最大地址是192.199.170.95。具体的计算方法是这样的：找出地址掩码中1和0的交界处发生在地址中的哪一个字节，现在是在第4个字节，因此只要把这一个字节的十进制82用二进制表示即可。十进制82的二进制是01010010，取其前3b(这3b加上前3个字节的24b等于前缀的27b)，再把后面5b都写成0，即01000000，等于十进制的64，这就找出了地址块的最小地址192.199.170.64。再把地址的第4个字节的最后5b都置1，即01011111，等于十进制的95，这就找出了地址块中的最大地址192.199.170.95。

3. 路由聚合(Route Aggregation)

CIDR可以把一些连续的多个小子网路由汇总成一条大网络的路由。这样通过这一条大网络路由条目就可以实现这些子网间的路由，最终实现减少路由表中的路由条目和提高路由性能的目的。CIDR可以看成是VLSM子网划分的逆过程，是把多个小子网汇聚成一个大的子网或标准有类网络。但必须同时借助VLSM来实现，所涉及的计算主要就是子网掩码，因为子网掩码与IPv4地址一起就可以确定路由中的具体目的网络。CIDR的路由聚合具有以下特点。

(1) 路由聚合有利于减少路由之间的路由选择信息交换，从而提升了整个因特网性能。

(2) 可以更有效地分配IPv4的地址空间。

(3) 路由表中的许多路由条目合并为更小的数目，这样减少路由器中路由表的大小，减少路由通告的时间。

其实路由聚合与前面介绍的子网划分实现机制上是一样的，就是通过改变网络的子网掩码长度来调整网络的大小。不同的只是子网划分是把网络ID向主机ID扩展(可以理解为向右走)，也就是网络ID向主机ID借位，缩小网络。而路由聚合则相反，是把主机ID向网络ID扩展(可以理解为向左走)，也就是主机ID向网络ID借位，扩大网络。当然，在实际的网络组建中，采用子网聚合的情况是极其少见的，CIDR路由汇总也是由路由器自动进行的。既然子网聚合和子网划分都是通过借位来实现的，所以它们的本质还是一样的。在子网划分中向主机ID借n位就可以划分成2^n个大小相等的连续子网；而在这里介绍的子网聚合中是向网络ID借n位就可以聚合2^n个大小相等的连续子网，即将网络前缀都相同的连续的IP地址组成“CIDR地址块”。

路由聚合的基本思想就是利用连续多个子网或标准网络的网络地址中相同的部分保留作为新网络的网络ID部分，不同的部分作为新网络的主机ID部分，然后由新网络的网络ID位数可以得出新网络的子网掩码；同时也可以求出聚合后新网络的地址范围、网络地址和广播地址。下面也通过几个具体的示例进行介绍。

例：聚合 192.168.4.0/24、192.168.5.0/24、192.168.6.0/24、192.168.7.0/24 这 4 个标准网络。

把这 4 个标准网络的网络地址转换成如下二进制形式：

11000000.10101000.00000100.00000000

11000000.10101000.00000101.00000000

11000000.10101000.00000110.00000000

11000000.10101000.00000111.00000000

从以上可以看出，4 个子网的网络地址中相同的部分就是用深颜色标注的这部分，相同部分共有 22b，即 11000000.10101000.000001，把这些位全部置 1，结果得出聚合后的子网掩码为 255.255.252.0。

例：聚合 192.168.5.0/24、192.168.6.0/24、192.168.7.0/24、192.168.8.0/24 这 4 个标准网络。

同样把这 4 个标准网络的网络地址转换成如下二进制形式：

11000000.10101000.00000101.00000000

11000000.10101000.00000110.00000000

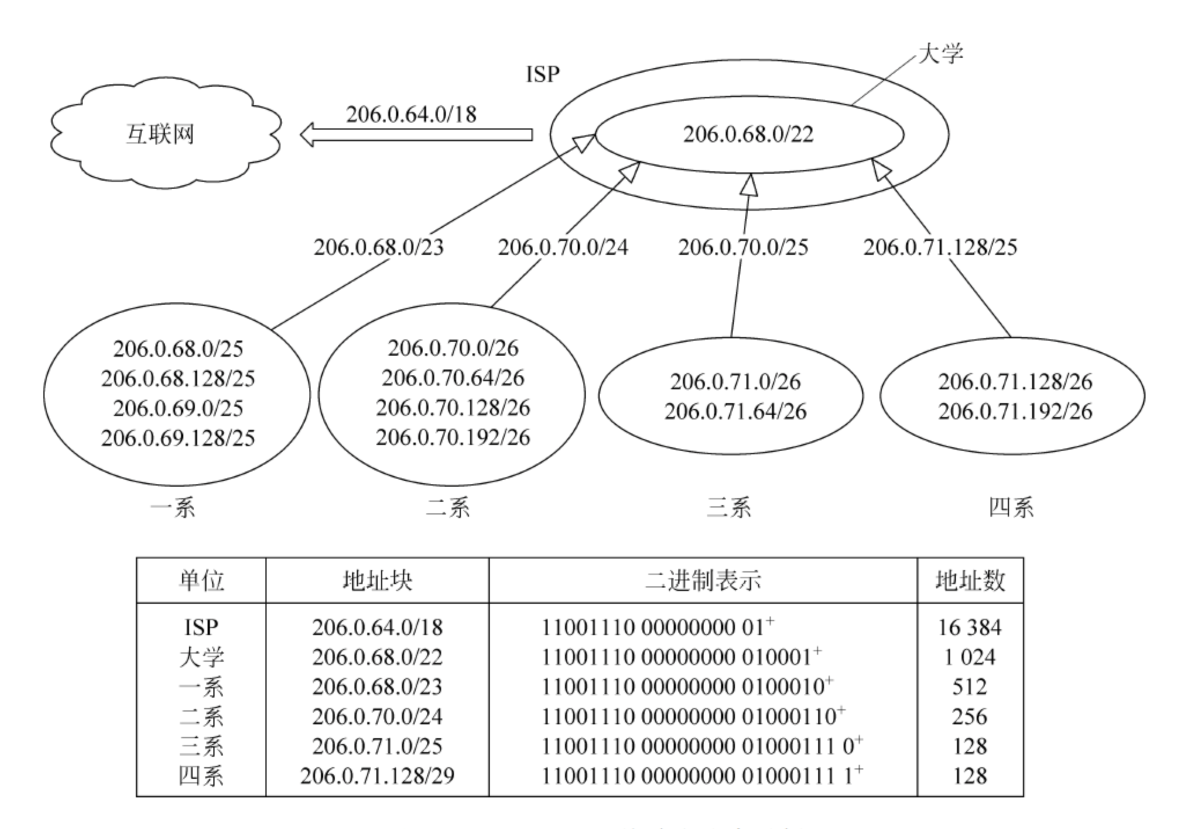
11000000.10101000.00000111.00000000

11000000.10101000.00001000.00000000

经过比较发现，这 4 个子网只有 20b 是完全相同的，也就是聚合后的子网掩码是 20b，而不是上例所得到的 22b，尽管它们聚合的都是 4 个连续的网路。究其原因是因为给出的这 4 个网络不能聚合成一个网络，而必须再结合前后连续的其他网络。其实本示例的聚合结果是对 192.168.0.0/24～192.168.15.0/24 共 16 个网络的聚合结果。

图 6.26 给出的是 CIDR 地址块分配的例子。假定某 ISP 已经拥有地址块 206.0.64.0/18，相当于拥有 64 个 C 类网络。现在某大学需要约 800 个 IP 地址，ISP 可以给该大学分配一个地址块 206.0.64.0/22，它包括 1024 个 IP 地址，相当于 4 个连续的 C 类/24 地址块，占该 ISP 拥有地址空间的 1/16。这所大学然后可以自由地对本校的各系分配地址块，而各系还可再划分地址块。CIDR 的地址块分配时不易看清，这主要是因为网络前缀和主机号的界限不是恰好出现在整数字节处，只要写出地址的二进制表示，并且主要将其中的一个关键字节转换成二进制的表示即可，弄清网络前缀的位数，就不会把地址块的范围弄错。

图 6.26 可以清楚看出地址聚合的概念，这个 ISP 共拥有 64 个 C 类网络。如果不采用 CIDR 技术，则在与该 ISP 的路由器交换路由信息的每一台路由器的路由表中，就需要 64 个项目。但采用地址聚合后，就只需用路由聚合后的一个项目 206.0.64.0/18 就能找到该 ISP。同理，这所大学共有 4 个系，在 ISP 内的路由器的路由表中，也需使用 206.0.64.0/22 这个项目。这个项目好比是大学收发室，凡是邮寄给这所大学任何一个系的邮件，邮递员都不考虑大学各个系的地址，而是把这些邮件投递到大学的收发室，然后大学收发室再进行下一步的投递，这样就减轻了邮递员的工作量，即简化了路由表的查找。从图 6.26 下面表格中的二进制可以看出，把 4 个系的路由聚合为大学的一个路由，即构成超网，是将网络前缀缩短。网络前缀越短，其地址块所包含的地址数就越多。而在三级结构的 IP 地址中，划分子网是使网络前缀变长。



单位	地址块	二进制表示	地址数
ISP	206.0.64.0/18	11001110 00000000 01+	16 384
大学	206.0.68.0/22	11001110 00000000 010001+	1 024
一系	206.0.68.0/23	11001110 00000000 0100010+	512
二系	206.0.70.0/24	11001110 00000000 01000110+	256
三系	206.0.71.0/25	11001110 00000000 01000111 0+	128
四系	206.0.71.128/29	11001110 00000000 01000111 1+	128

图 6.26　CIDR 地址块路由聚合示例

6.5　拥塞控制

6.5.1　拥塞的概念

当正在通过网络传输的分组的数目开始接近网络的分组处理能力时，网络就会来不及处理这些分组，以致引起部分乃至整个网络性能下降的现象，这称为拥塞，如图 6.27 所示。拥塞产生的后果如图 6.28 所示。

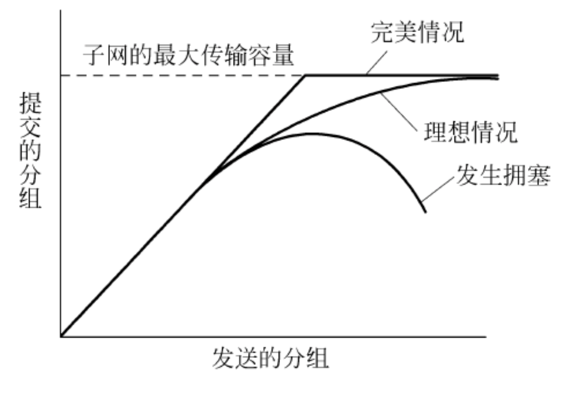


图 6.27　拥塞的概念

造成拥塞有多种原因。短时间内分组流突然从多个输入到达，且要输出到同一条线路，这时就需要有存储器暂存，并建立队列。如果此时没有足够的缓存空间来保存这些分组，有些分组就会丢失。处理器速度慢也能导致拥塞。在由路由器互联的网络中，如果路由器的 CPU 处理速度太慢，以至于不能执行日常工作，那么，即使有足够的线路容量，也可能出现队列饱和。类似地，低带宽线路也会导致拥塞。

拥塞可能导致恶性循环。如果路由器中没有空余的缓冲区，则必须丢掉新到的分组。分组被丢掉，发送方会因为超时而重发此分组，而且可能要重发多次。由于发送方在未接到

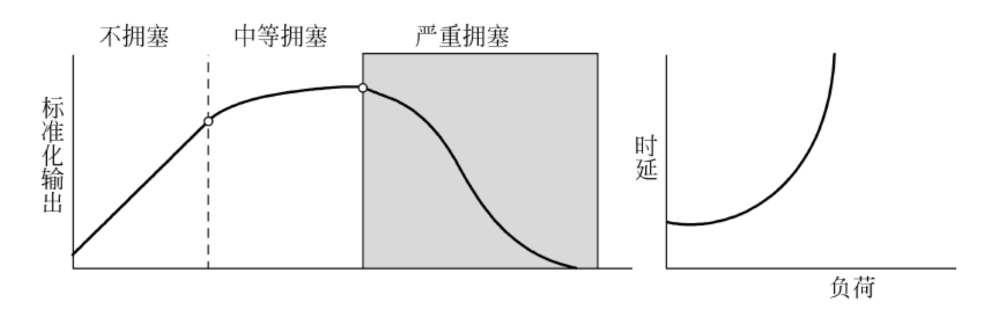


图 6.28 拥塞产生的后果

确认之前不能扔掉该分组,因此接收端的拥塞也导致了发送方缓冲器得不到释放。这样就使拥塞加重了。

需要指出的是,拥塞控制和流量控制的差异。拥塞控制必须确保通信子网能运送待传输的数据,这是全局性的问题,它涉及所有主机、路由器,路由器中存储转发处理的行为,以及所有导致削弱通信子网负荷能力的其他因素。而流量控制只与某发送者和某接收者之间的点到点通信量有关。它的任务是确保一个快速发送者不能以比接收者能承受的速率更高的速度传输数据。流量控制总是涉及接收者,接收者要向发送者送回另一端情况的一些直接反馈。

拥塞控制与流量控制容易混淆的原因在于,有些拥塞控制算法在网络出现问题时,通过往各源端发送消息来告诉它们要减慢发送速度。

6.5.2 拥塞控制的基本原理

解决拥塞控制的方案被分为两类:一类是开环;另一类是闭环。开环的关键在于,它致力于通过良好的设计来避免问题的出现,确保问题在一开始就不会发生。一旦系统安装并运行起来,就不再做任何中间阶段的更正了。而闭环的解决方案是建立在反馈环路的概念上的。建立在反馈环路基础上的闭环解决拥塞控制方案,分为以下 3 个部分。

(1) 检测系统何时何地发生了拥塞。度量子网拥塞状况的方式主要包括:因缺少缓冲区空间而丢失分组的比例,平均队列长度,超时和重发分组的数量,平均分组延迟等。对于以上所有因素,数值的增加就是发生拥塞可能性的增加。

(2) 将拥塞信息从检测点传送到可能对此采取行动的地方。最简便的方法就是检测到拥塞的路由器向信息源发送一个分组。但是,这些额外的分组又给已经拥塞的子网增加了负荷。也可以在每个分组中保留一位或一个字段,当路由器发现拥塞状态超过某个临界值时,就在所有发送的分组中填充这个位或字段,以警告它的“邻居”要发生拥塞了。

(3) 调整系统操作来修正问题。

6.5.3 拥塞控制的方法

在分组交换网络、帧中继网络、ATM 网络以及基于 IP 的互联网中,用于拥塞控制的技术各种各样,常用的技术有反压、阻流分组、隐式拥塞信令和显式拥塞信令 4 种,如图 6.29 所示。

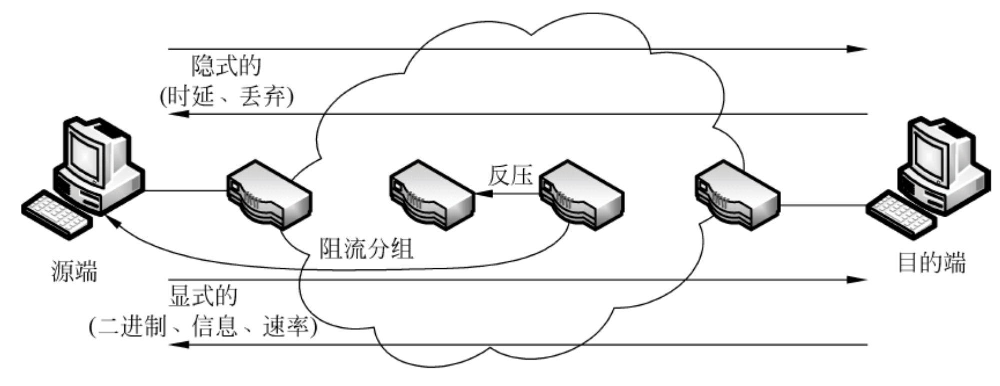


图 6.29 拥塞的控制机制

1. 反压

反压(Backpressure)技术产生的效果类似于流过管道的流体产生的反压现象。当管道末端被关闭(或受限制)时,流体就向源头产生压力,从而阻断(或减慢)流量。

反压可在链路或逻辑连接(如虚电路)的基础上实施。如图 6.30 所示,如果节点 6 变得拥挤(缓存满溢),那么节点 6 就可以减慢或阻止来自节点 5(或节点 3,或节点 5 和节点 3)的所有分组流量。如果这一限制持续下去,节点 5 会减慢或阻止本身的入口链路上的通信量。这种流量的限制将会反向(与数据通信流反方向)传播到信源,从而信源限制新的分组流入网络。

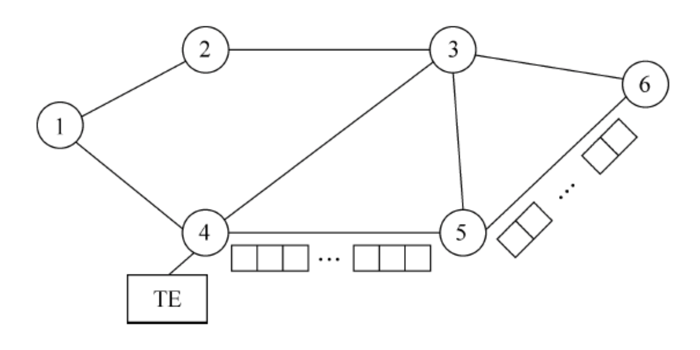


图 6.30 数据网络中队列与队列的相互作用

反压可以有选择地应用到某些逻辑连接上,以便从一个节点到下一个节点之间的流量只在某些连接上受限或停止,它们通常是通信量最大的连接。在这种情况下,流量的限制就会沿着该连接传播到数据源。

反压的使用比较有限。它可以用在允许使用逐跳(从一个节点到下一个节点)流量控制的面向连接的网络,基于 X.25 的分组交换网通常支持这一功能。但帧中继和 ATM 网络都没有逐跳限制流量的功能。对于基于 IP 的互联网,它的内部也没有设施可以调整那些沿着互联网通路从一台路由器到下一台路由器的数据流量。

2. 阻流分组

阻流分组(Choke Packet)是拥塞的节点产生的控制分组,并且被传回源节点以限制通信流量。Internet 控制报文协议(ICMP)的“源抑制(Source Quench)”分组就是阻流分组的例子。路由器或目的端系统都可以向源端系统发送此报文,以要求源端系统减小对目的站的发送速度。源站接收到一个抑制报文,主机就会降低对特定的目的站的发送速率,直至不再收到源站抑制报文。当路由器或主机由于缓存溢出而不得不丢弃 IP 数据报时,就可以使用源站抑制报文,并且会为每个被丢弃的数据报发送一个抑制报文。另外,当系统的缓存即

将存满时，有可能预测到拥塞的发生，并发送源站抑制报文；这时，抑制报文所指的那个数据报可能已经正常地被传递到目的地了。因此，接收到源站抑制报文并不表示相应的数据报被正确传递。

3. 隐式拥塞信令

网络在发生拥塞时，可能会出现两种可能：①从源站到目的站的个别分组的传输时延加大，以至于明显比固定传播时延长得多；②分组被丢弃。如果源站能够检测到传输时延的增加或者分组的丢弃，就说明网络发生了拥塞。如果所有源站都能检测到拥塞，并因此减缓流量，网络拥塞就可以消除了。因此，基于隐式拥塞信令的拥塞控制是由端系统完成的，并不需要网络节点的参与。

在像数据报分组交换网和基于 IP 的互联网这样的无连接或数据报配置中，隐式拥塞信令是一种有效的拥塞控制技术。在这种情况下，整个互联网层并没有能够调整流量的逻辑连接，但在两个端系统之间，可以在 TCP 层建立逻辑连接。TCP 包括对收到的报文进行确认的机制，以及以源站到目的站的 TCP 连接为单位的管理数据流量的机制。

4. 显式拥塞信令

显式拥塞信令技术所要实现的目标是希望网络中的可用容量都能够得到充分利用，但同时又要能够对拥塞做出及时的反应从而控制拥塞。

在显式拥塞信令技术中，网络会对正在形成的拥塞向端系统发出警告，而端系统则应采取措施降低对网络的供给负荷。显式拥塞信令技术通常用于面向连接的网络中，并以单个连接为单位控制分组流量。显式拥塞信令可以向如下两个方向发送。

(1) 向反(Backward)：通知源站点应该对收到的分组方向相反的通信量采取必要的拥塞控制措施。它表示用户在该逻辑连接上传输的分组遇到了拥塞的网络资源。向反信息的传送是通过源站数据分组中的某些改变了的比特，或者是给源站单独发送的控制分组。

(2) 向前(Forward)：通知用户应该对与收到的分组方向相同的通信量采取必要的拥塞控制措施。它表示这个分组在其逻辑连接上遇到了拥塞的网络资源。同样，这个分组也可以通过数据分组中改变了的比特，或者单独的控制分组来传递。在某些方案中，端系统收到向前信号时，会将此信号沿着同一条逻辑连接返回给源端；在另一些方案中，端系统将在高层(如 TCP)对源端系统实施流量控制措施。

6.5.4 通信量整形

拥塞发生的主要原因在于通信量往往是突发性的。如果主机能够以一个恒定的速率发送信息，则拥塞就会少很多。有一种管理拥塞的方法是强迫分组以某种更有预见性的速率发送，称为通信量整形(Traffic Control Shaping)。通信量整形是调整数据传输大的平均速率(以及突发性)，而滑动窗口协议只是限制一次能传送数据的量，而不是传送的速率。

监视一个通信量称为通信量控制策略(Traffic Control Policing)。常用的通信量整形算法是漏桶算法(Leaky Bucket Algorithm)和令牌桶算法(Token Bucket Algorithm)。假设一只底部有一个小孔的桶，不管进水的速率多大，水从桶中往外漏的速率是恒定的，一旦桶空了，速率便为 0，而且只要桶满了，再往里流的水都会从桶边溢出去流失了。

漏桶算法用于分组，如图 6.31 所示。概念上每台主机都通过一个包含漏桶的接口与网络相连，漏桶是一个有限的内部队列，如果分组到达队列时队列已满，分组就会被丢弃。也就是

说，当队列的长度已达到最大值时，如果主机中的一个或者多个进程仍然试着想发送分组，这些分组将会被丢掉。这种算法被称为漏桶算法，它实际上只不过是一个有恒定服务时间的单服务器队列系统。应用原始漏桶算法很简单，漏桶由一个有限队列构成。当分组到达时，如果队列未满，则将其加到队尾；否则丢弃它。每个时钟节拍发送一个分组(除非队列为空)。

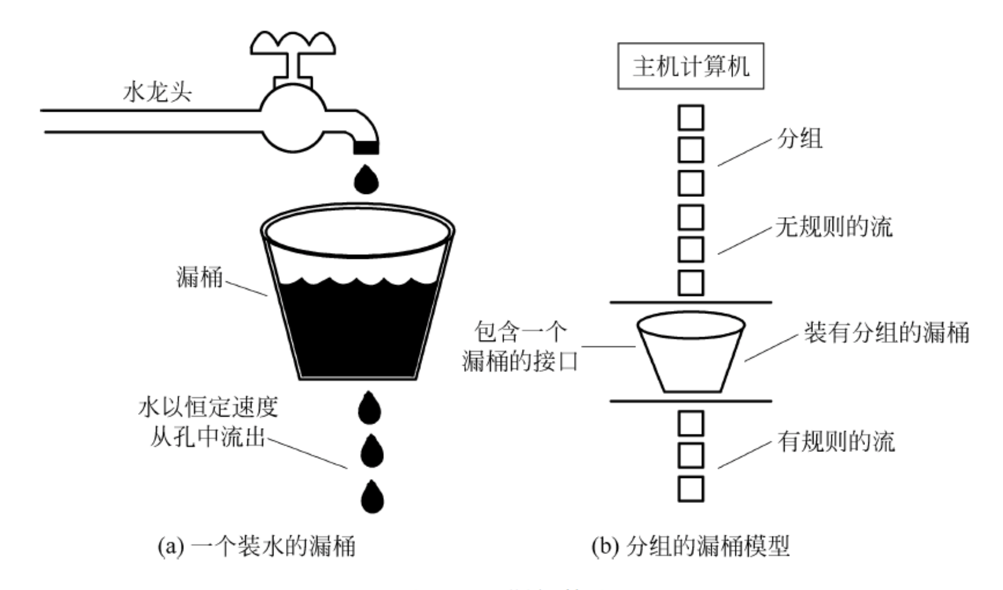


图 6.31　**漏桶算法**

令牌桶算法是指当大的突发通信量到来时，输出也相应加速一点，也是一个更有弹性的、更合适的绝不会丢失数据的算法。而漏桶算法强迫输出模式保持一个固定的平均速率，而不管突发通信量的大小。在此算法中，漏桶可以保留令牌，由一个时钟每隔 ts 生成一个令牌，如图 6.32(a)所示。图中有一个保留着 3 个令牌的桶，有 5 个分组等着传送，每传送一个分组就必须得到和消耗一个令牌。在图 6.32(b)中，有 5 个分组中的 3 个已经被传送出去，但其余 2 个还在等待新令牌的生成。

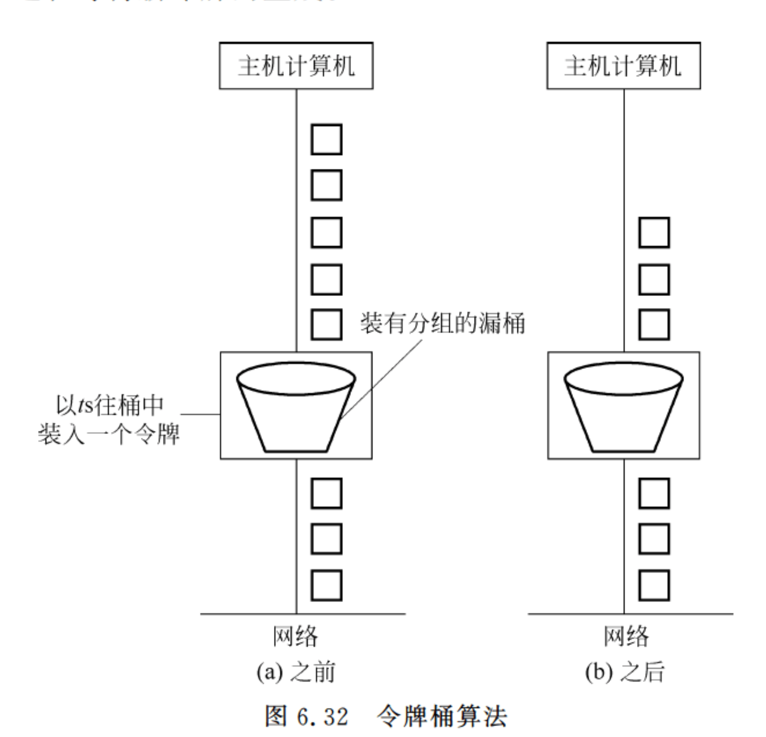


图 6.32　**令牌桶算法**

6.6 下一代互联网的网际协议 IPv6

6.6.1 IPv6 的基本概念

现有互联网是在 IPv4 协议的基础上运行。随着互联网的迅速发展，IPv4 定义的有限地址空间逐渐被耗尽，地址空间的不足必将影响互联网的进一步发展。IPv4 采用 32b 地址长度，只有大约 43 亿个地址，现已被分配完毕。

IPv6 是 Internet Protocol Version 6 的缩写，其中 Internet Protocol 译为"互联网协议"。IPv6 是 IETF(Internet Engineering Task Force，互联网工程任务组)设计的用于替代现行版本 IP 协议(IPv4)的下一代 IP 协议。IPv6 采用 128b 地址长度，按保守方法估算 IPv6 实际可分配的地址，整个地球每平方米面积上可分配 1000 多个地址。在 IPv6 的设计过程中还考虑了在 IPv4 中解决不好的其他问题。

如果说 IPv4 实现的只是人机对话，而 IPv6 则扩展到任意事物之间的对话，它不仅可以为人类服务，还将服务于众多硬件设备，如家用电器、传感器、远程照相机、汽车等，它将是无时不在、无处不在地深入社会每个角落的真正的宽带网。而且它所带来的经济效益将非常巨大。

与 IPv4 相比，IPv6 具有以下几个优势。

(1) IPv6 具有更大的地址空间。IPv4 中规定 IP 地址长度为 32b，即有 $2^{32}-1$ 个地址；而 IPv6 中 IP 地址的长度为 128b，即有 $2^{128}-1$ 个地址。

(2) IPv6 使用更小的路由表。IPv6 的地址分配一开始就遵循聚类(Aggregation)的原则，这使得路由器能在路由表中用一条记录(Entry)表示一片子网，大大减小了路由器中路由表的长度，提高了路由器转发数据包的速度。

(3) IPv6 增加了增强的组播(Multicast)支持以及对流的支持(Flow Control)，这给网络上的多媒体应用提供了长足发展的机会，为服务质量(Quality of Service，QoS)控制提供了良好的网络平台。

(4) IPv6 加入了对自动配置(Auto Configuration)的支持。这是对 DHCP 协议的改进和扩展，使得网络(尤其是局域网)的管理更加方便和快捷。

(5) IPv6 具有更高的安全性。在使用 IPv6 网络中用户可以对网络层的数据进行加密并对 IP 报文进行校验，极大地增强了网络的安全性。

在 IPv6 中，可使用以下 3 种形式表示 IP 地址。

(1) 冒号十六进制形式。这是首选形式 n:n:n:n:n:n:n:n，由 8 个 16b 地址元素组成，每个地址元素用十六进制值表示。例如，3FFE:FFFF:7654:FEDA:1245:BA98:3210:4562。

(2) 压缩形式。由于地址长度要求，地址中经常包含由 0 组成的长字符串。为了简化对这些地址的写入，可以使用压缩形式。在这一压缩形式中，多个 0 块的单个连续序列由双冒号符号(::)表示。此符号只能在地址中出现一次。例如，多路广播地址 FFED:0:0:0:0:BA98:3210:4562 的压缩形式为 FFED::BA98:3210:4562。单播地址 3FFE:FFFF:0:0:8:

800:20C4:0 的压缩形式为 3FFE:FFFF::8:800:20C4:0。环回地址 0:0:0:0:0:0:0:1 的压缩形式为::1。未指定的地址 0:0:0:0:0:0:0:0 的压缩形式为::。

(3) 混合形式。此形式组合 IPv4 和 IPv6 地址。在此情况下,地址格式为 n:n:n:n:n:n:d.d.d.d,其中每个 *n* 都表示 6 个 IPv6 高序位 16b 地址元素之一的十六进制值,每个 *d* 都表示 IPv4 地址的十进制值。

6.6.2 由 IPv4 向 IPv6 过渡

美国国防部因为 IPv6 在安全方面的优势而大力支持它。其在 2003 年 7 月宣布,在 2008 财年,将全部完成向 IPv6 的转移;同年 10 月宣布,所有开发、生产、采购的网络资产都必须过渡到 IPv6。

在向 IPv6 转移方面,中国走到了前列。2004 年 12 月 17 日,中国第一个下一代 Internet 暨中国下一代 Internet 示范工程核心网(CERNET2)正式开通,这是世界上规模最大的纯 IPv6 Internet。

对于商业用户和传统的消费者,连接到网上的设备和应用的数量是不可预测的。IPv6 扩大了家庭网络的应用空间,包括应用管理、多媒体娱乐设备和家庭安全。这些应用特别是家庭安防设施需要端到端的身份识别与数据加密。拥有 IPv6,DSL 和 Cable Modem 用户可以建立自己的家庭网络,可以远程安全地监视和控制家庭网络设备。

为了实现从 IPv4 向 IPv6 的过渡,人们已经根据不同的应用情况,设计了多种过渡技术和解决方案,大致分为以下 3 类。

(1) IPv4/IPv6 双栈技术。IPv4/IPv6 双栈技术是最主要的过渡机制。在网络一侧的接入服务器上实现双栈,成为 IPv4 与 IPv6 的接入点,使终端接入 IPv4 与 IPv6 服务,以免在网络中使用额外的翻译器。

(2) 隧道技术。隧道技术在一端把 IPv6 包封装为 IPv4 包的数据内容,然后在另一端解封复原成 IPv6 包。隧道要求在封装/解封的节点上有 IPv4/IPv6 双栈能力。这可能是未来采用最多的一种方式。

(3) 翻译器技术。翻译器是一个处在纯 IPv4 终端和纯 IPv6 终端之间的部件,它可使这些终端之间能直接进行通信,且不需要对终端进行任何修改。

6.6.3 IPv6 地址方案

和 IPv4 相比,IPv6 的主要改变就是地址的长度为 128b,也就是说,可以有 $2^{128}-1$ 个 IP 地址,相当于 10 的后面有 38 个 0,足以保证地球上的每个人拥有一个或多个 IP 地址。

1. IPv6 地址类型

在 RFC1884 中指出了 3 种类型的 IPv6 地址,它们分别占用不同的地址空间。

(1) 单播:单一接口的地址。发送到单播地址的数据报被送到由该地址标识的接口。

(2) 任意播放:一组接口的地址。大多数情况下,这些接口属于不同的节点。发送到任意播放地址的数据报被送到由该地址标识的其中一个接口。由于使用任意播放地址的标准尚在不断完善中,所以目前 HP-UX 不支持任意播放。

（3）多播：一组接口的地址（通常分属不同节点）。发送到多播地址的数据报被送到由该地址标识的每个接口。

和 IPv4 不同的是，IPv6 中出现了任意点传输地址，并以多点传输地址代替了 IPv4 中的广播地址。

2. IPv6 地址分配

RFC1881 规定，IPv6 地址空间的管理必须符合 Internet 团体的利益，必须通过一个中心权威机构来分配。目前，这个权威机构就是 IANA（Internet Assigned Numbers Authority，Internet 分配号码权威机构）。IANA 会根据 IAB（Internet Architecture Board）和 IEGS 的建议来进行 IPv6 地址的分配。

目前，IANA 已经委派下述 3 个地方组织来执行 IPv6 地址分配的任务。

（1）欧洲的 RIPE-NCC（www. ripe. net）。

（2）北美的 INTERNIC（www. internic. net）。

（3）亚太地区的 APNIC（www. apnic. net）。

6.6.4 IPv6 地址表示方法

现有的 IP 地址（IPv4 IP 地址）是用 4 段十进制数的数字，用“.”隔开来表示，每一段如用二进制数表示则包含 8b。IPv6 的地址在表示和书写时，用冒号将 128b 分割成 8 个 16b 的段，这里的 128b 表示在一个 IPv6 地址中包括 128 个二进制数。

1. IPv6 地址的文本表示

有 3 种常规格式可用于以文本字符串形式表示 IPv6 地址。

（1）x:x:x:x:x:x:x:x，其中，x 是十六进制数值，分别对应于 128b 地址中的 8 个 16b 区段。例如，2001:fecd:ba23:cd1f:dcb1:1010:9234:4088。

（2）一些 IPv6 地址可能包含一长串 0 位。为了便于以文本方式描述这种地址，制定了一种特殊的语法。::的使用表示有多组 16b 零。::只能在一个地址中出现一次，可用于压缩一个地址中的前导、末尾或相邻的 16b0。例如，fec0:1:0:0:0:0:0:1234 可以表示为 fec0:1::1234。

（3）当处理拥有 IPv4 和 IPv6 节点的混合环境时，可以使用 IPv6 地址的另一种形式，即 x:x:x:x:x:x:d. d. d. d。其中，x 是 IPv6 地址的 96b 高位顺序字节的十六进制数值，d 是 32b 低位顺序字节的十进制数值。通常，“映射 IPv4 的 IPv6 地址”以及“兼容 IPv4 的 IPv6 地址”可以采用这种表示法表示。例如，0:0:0:0:0:0:10. 1. 2. 3 以及::10. 11. 3. 123。

2. IPv6 地址前缀

IPv6 地址前缀与 IPv4 中的 CIDR 相似，并写入 CIDR 表示法中。IPv6 地址前缀表示为 IPv6-address/prefix-length。其中，IPv6-address 是用上面任意一种表示法表示的 IPv6 地址，prefix-length 是一个十进制数值，表示前缀由多少个最左侧相邻位构成。例如，fec0:0:0:1::1234/64。地址的前 64bfec0:0:0:1 构成了地址的前缀。在 IPv6 地址中，地址前缀用于表示 IPv6 地址中有多少位表示子网。

3. 单播地址

IPv6 单播地址分为多种类型，分别是全局可聚集单播地址、站点本地地址以及链路本

地地址。通常,单播地址在逻辑上如图 6.33 所示。

n 位	128−*n* 位
子网前缀	接口 ID

图 6.33 单播地址逻辑结构

IPv6 单播地址中的接口标识符用于在链路中标识接口。接口标识符在该链路中必须是唯一的,链路通常由子网前缀标识。

如果一个单播地址的所有位均为 0,那么该地址称为未指定的地址。以文本形式表示为::。

单播地址::1 或 0:0:0:0:0:0:0:1 称为环回地址。节点向自己发送数据报时采用环回地址。

6.6.5 IPv6 数据报格式

IPv6 数据报格式由 3 部分组成:IPv6 数据报报头、扩展(下一个头标)和高层数据。IPv6 数据报报头格式用图 6.34 来表示,各项具体的含义可以通过表 6.8 进行说明。

版本	优先级	流标记	有效载荷长度	下一个头标	站段限制	源地址	目的地址
4	4	24	16	8	8	128	128

共320b

图 6.34 IPv6 数据报报头格式

表 6.8 IPv6 数据报报头各项作用

IPv6 数据报报头项	作　　用
版本(Version)	IPv6 协议中规定该字段值为 6
优先级(Priority)	当该字段为 0~7 时,表示在拥塞发生时允许进行延迟处理,值越大优先级越高;当该字段为 8~15 时,表示处理以固定速率传输的实时业务,值越大优先级越高
流标记(Flow Label)	路由器根据流标记的值在连接前采取不同的策略
有效载荷长度(The Next Header)	指扣除报头后的净负载长度
下一个头标(The Next Header)	如果该数据有附加的扩展头,则该字段标识紧跟的下一个扩展头;若无,则标识传输层协议种类,如 UDP(17)、TCP(6)
站段限制(Hop Limit)	即转发上限,该字段是为防止数据报传输过程中无休止地循环下去而设定的。该项首先被初始化,然后每经过一台路由器该值就减 1,当减为 0 仍未到达目的端时就丢弃该数据报
源地址(Source Address)	发送方 IP 地址,128b
目的地址(Destination Address)	接收方 IP 地址,128b

6.6.6 从 IPv4 到 IPv6 的过渡

尽管 IPv6 比 IPv4 具有明显的先进性,但要在短时间内将 Internet 和各个企业网络中的所有系统全部从 IPv4 升级到 IPv6 是不可能的,IPv4 的网络将在相当长时间内和 IPv6 的网络共存。为了促进和保证 IPv4 网络向 IPv6 网络的平滑迁移,IETF 专门成立了 NGtrans 工作组,以负责制定 IPv4 向 IPv6 过渡的方案。该工作组所制定的过渡机制包括双栈过渡机制、基于隧道的过渡机制和基于协议转换的过渡机制等,这些不同的过渡机制分别适用不同的场合。

1. 双栈过渡机制

双协议栈过渡机制简称双栈,是一种最直接的过渡机制。该机制在网元(注:包括主机和路由器)的 IP 层同时实现 IPv4 和 IPv6 两种协议。由于同时实现了 IPv4 和 IPv6 协议,因此各网元在通过 IPv4 协议与现有的 IPv4 网络通信的同时,可以通过 IPv6 协议与新建的 IPv6 网络通信。

图 6.35 给出了一个双栈网元中,高层的应用使用协议栈的情况。当主机或者路由器提供双栈协议之后,原有的不支持 IPv6 协议的 IPv4 应用可以继续使用 IPv4 协议栈来与其他节点通信。而那些支持 IPv6 的新应用一般同时也兼容 IPv4,因此在利用网络层的 IP 协议栈与其他节点通信时,就可以根据 DNS 解析的结果,选择使用 IPv4 或者 IPv6 协议栈。

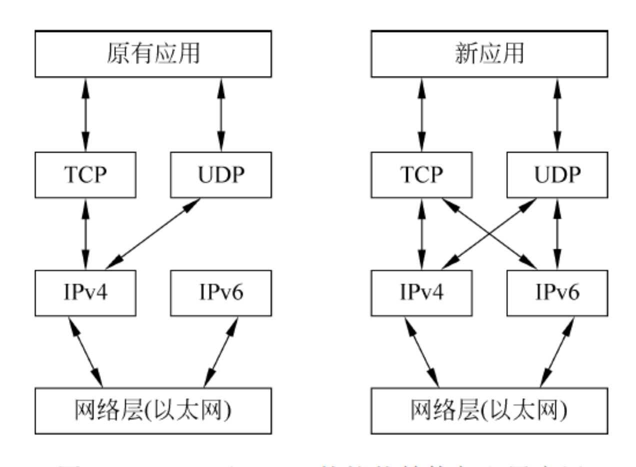


图 6.35 IPv4/IPv6 双协议栈结构与上层应用

尽管双协议栈是实现 IPv4 和 IPv6 兼容的一种最为直接的方法。但是由于需要同时支持 IPv4 和 IPv6 两种协议,因此整个协议栈的结构比较复杂。特别是对于双栈路由器,不仅需要同时运行 IPv4 下的路由协议和 IPv6 下的路由协议,同时还需要保存两套分别针对 IPv4 和 IPv6 的路由表,从而要求路由器提供较高的 CPU 处理能力和更多的内存资源。如果将双栈过渡机制用于骨干网,则需要对大量的网络设备进行升级,其难度比较大。因此,在现阶段双栈网元一般只用于 IPv4 网络或者 IPv6 网络的边缘,作为隧道过渡机制的隧道端点部署,以解决 IPv4 或者 IPv6 网络直接的互通问题。

2. 隧道过渡机制

在 IPv6 开始部署的早期阶段,IPv6 网络相对于已有的 IPv4 互联网就像海洋中的孤岛。这些没有直接连接的 IPv6 孤岛被 IPv4 海洋分隔开来,为了在这些 IPv6 孤岛之间进行

通信，就必须保证 IPv6 报文能够从一个 IPv6 网络出发，穿过 IPv4 互联网，到达目的端的 IPv6 网络。隧道过渡机制就是解决该问题的一个比较直接的方法。

所谓“隧道”就是在 IPv6 网络和 IPv4 网络邻接的双栈路由器上，利用 IPv4 报文封装 IPv6 报文，然后完全按照 IPv4 的路由策略将该报文发送到接收端网络中与目的 IPv6 网络邻接的另外一个双栈路由器，由该路由器将封装在 IPv4 报文中的 IPv6 报文解封装，然后利用 IPv6 的路由策略完成 IPv6 报文的最终转发和处理的过程。IPv4 隧道就是一个虚拟的点到点连接，对于其所连接的 IPv6 网络或者所通过的 IPv4 网络来说都是透明的，只需对隧道的起点和终点进行升级即可。图 6.36 给出了一个利用 IPv4 隧道实现 IPv6 网络互联的例子。

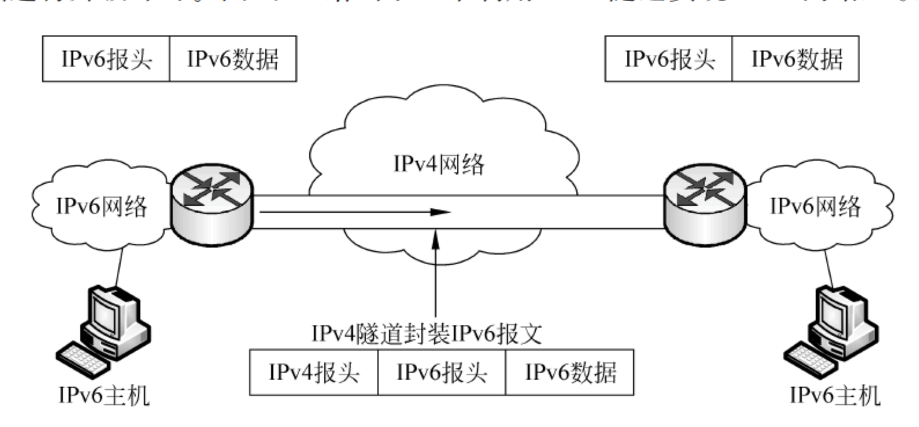


图 6.36　利用 IPv4 隧道实现 IPv6 网络互联

3. NAT-PT

隧道方式一般用于源与目的均为 IPv6 网络的互联互通环境，当 IPv6 网络中不支持 IPv4 的节点需要和 IPv4 网络中不支持 IPv6 的节点进行通信时，隧道方式就不再适用了，此时需要使用协议转换的方法。网络地址翻译－协议转换（Network Address Translation-Protocol Translation，NAT-PT）技术就是一种利用协议转换来实现纯 IPv6 网络和纯 IPv4 网络之间互通的方法。

图 6.37 给出了使用 NAT-PT 进行 IPv6 和 IPv4 网络互通的简单示意。当右边的 IPv6 网络需要与左边的 IPv4 网络相互通信时，不得不通过位于它们之间的 NAT-PT 转换网关对报文的地址和格式等信息进行必要的转换，以实现两种不同类型 IP 网络的互联。

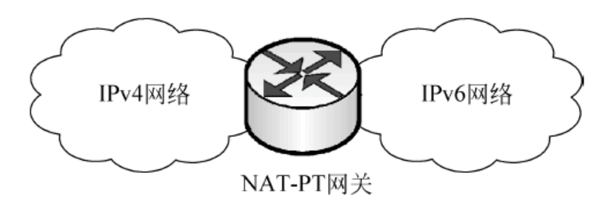


图 6.37　利用 NAT-PT 进行 IPv6 和 IPv4 的互联

6.7　移动 IP

移动 IP 技术是指移动用户可跨网络随意移动和漫游，不用修改计算机原来的 IP 地址，同时，继续享有原网络中一切权限。简单地说，移动 IP 就是实现网络全方位的移动或者漫

游。IETF 为了满足这种需求，制定了移动 IP 协议，从而使因特网上的移动接入成为可能。目前，IETF 正在开发一套用于移动 IP 的技术规范，这主要是 RFC 2002、RFC 2003、RFC 2004、RFC 2290。

6.7.1 移动 IP 的出现

因特网的飞速发展和移动计算通信设备（便携计算机、PDA 等）日益广泛的应用，推动了无线接入的研究和移动因特网的研究。移动计算机用户希望接入同样的网络，共享资源和服务，而不局限于某一固定区域。当它移动时，也能方便地断开原来的连接，并建立新的连接。如果节点从一条链路切换到另一条链路而没有改变它的 IP 地址，那么它就不可能在新链路上接收到数据包。

解决方案如下。

(1) 根据主机地址进行路由选择。这种方法将大量浪费路由器的有限资源，显然不能满足网络互联的要求。

(2) 在移动节点每次变换位置时，改变其 IP 地址。这种方法需频繁更新域名系统(DNS)服务器，所以也不可取。

(3) 在数据链路层使用蜂窝数字分组数据(CDPD)标准。这种方法需要新的网络基础设施和大量管理维护费用，且无法与现存的国际互联网兼容，也不是合适的解决方案。

移动 IP 技术引用了处理蜂窝移动电话呼叫的原理，使移动节点采用固定不变的 IP 地址，一次登录即可实现在任意位置上保持连接，使通信持续进行。

6.7.2 移动 IP 的基本术语

本节介绍与移动 IP 技术相关的重要术语，如图 6.38 所示。

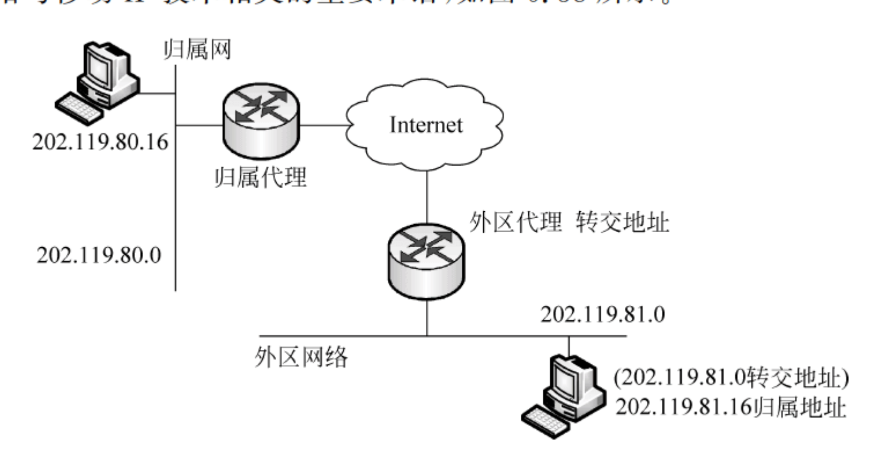


图 6.38 术语图示

1. 移动代理

移动代理(Mobility Agent)分为归属代理(Home Agent)和外区代理(Foreign Agent)两类，它们是服务器或路由器，能知道移动节点实际连接在何处。

归属代理又称为家乡代理，是一个在移动节点归属网(Home Network)上的路由器，它

至少有一个接口在归属网上，当移动节点离开归属网后，它通过IP隧道把数据包转发给移动节点，并且负责维护移动节点的当前位置信息。

外区代理位于移动节点当前连接的外区网络上，它向已注册的移动节点提供路由服务。当使用外区代理转交地址时，外区代理负责解除原始数据包的隧道封装，取出原始数据包，并将其转发到该移动节点。对于那些由移动节点发出的数据包而言，外区代理可作为已注册的移动节点的默认路由器使用。

2. 移动IP地址

移动IP节点拥有如下两个IP地址。

1）归属地址

归属地址（Home Address）又称为本地地址，这是用来识别端到端连接的静态地址，也是移动节点与归属网连接时使用的地址。不管移动节点连至网络何处，其归属地址保持不变。

2）转交地址

转交地址就是隧道终点地址，转交地址可能是外区代理转交地址，也可能是驻留本地的转交地址（配置转交地址）。通常用的是外区代理转交地址。在这种地址模式中，外区代理就是隧道的终点，它接收隧道数据包，解除数据包的隧道封装，然后将原始数据包转发到移动节点。

转交地址是一个临时分配给移动节点的地址。它由外部获得（如通过DHCP），移动节点将其与自身的一个网络接口相关联。一个配置转交地址仅能被一个移动节点使用。当使用驻留归属的转交地址时，移动节点自身就是隧道的终点，执行解除隧道功能。

转交地址是仅供数据包路由使用的动态地址，也是移动节点与外区网连接时使用的临时地址。每当移动节点接入一个新的网络，转交地址就发生变化。

3. 位置注册

移动节点必须将其位置信息向其归属代理进行注册（Registration），以便被找到。在移动IP技术中，按照网络连接方式的不同，有如下两种不同的注册规则（图6.39）。

（1）通过外区代理进行注册，即移动节点向外区代理发送注册请求报文，外区代理接收并处理注册请求报文，然后将报文中继到移动节点的归属代理。

（2）直接向归属代理进行注册，即移动节点向其归属代理发送注册请求报文，归属代理处理后向移动节点发送注册答复报文。

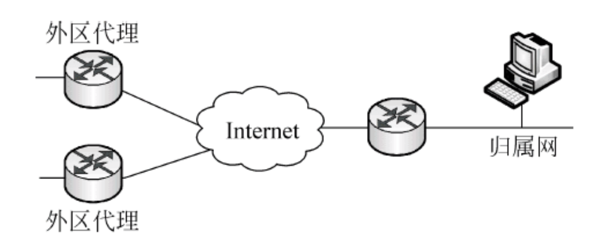


图6.39　实体之间的联系

4. 代理发现

为了随时随地与其他节点进行通信，移动节点必须实现代理发现（Agent Discovery）。移动IP定义了两种发现移动代理的方法：一个是被动发现，移动节点等待代理周期性地广

播代理通告报文；另一个是主动发现，移动节点广播一条请求代理的报文。

所有移动代理都应具备代理通告功能，并对代理请求做出响应。所有移动节点必须具备代理请求功能。但是，移动节点只有在没有收到移动代理的代理通告，并且无法通过数据链路层协议或其他方法获得转交地址的情况下，方可发送代理请求报文。

5. 隧道技术

当移动节点在外区网上时，归属代理需要将原始数据报转发给已注册的外区代理。此时，归属代理使用 IP 隧道技术(Tunneling)，将原始 IP 数据报封装在转发的 IP 数据报中，从而使原始 IP 数据报原封不动地转发到处于隧道终点的转交地址处。

在转交地址处解除隧道，取出原始数据报，并将原始数据报发送到移动节点。当转交地址为配置转交地址时，移动节点本身就是隧道的终点，它自身进行解除隧道，取出原始数据报的工作。

RFC 2003 和 RFC 2004 中分别定义了两种隧道封装技术，如图 6.40 所示。

用 IP 封装 IP[RFC 2003]，需要在原始数据包的现有首部前插入一个外层 IP 首部。外层 IP 首部中的源地址和目的地址分别标识隧道的两个边界节点。

最小封装[RFC 2004]，数据包在封装之前不能被分片。对移动 IP 技术来讲，最小封装技术是可选的。为了使用最小封装技术来封装数据包，移动 IP 技术需要在原始数据包经修改的 IP 首部和未修改的净负荷之间插入最小转发首部。

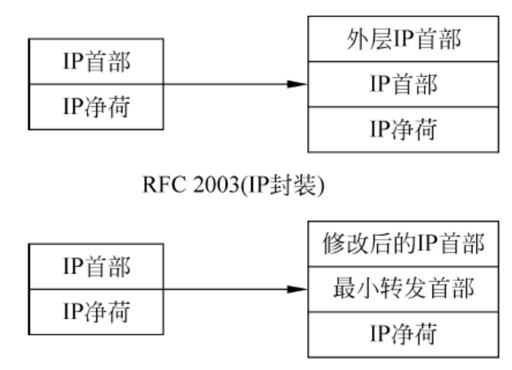


图 6.40 两种隧道封装技术

6.7.3 移动 IP 的工作原理

移动 IP 的工作原理由以下步骤实现。

(1) 移动 IP 系统中的归属代理和外区代理不停地向网上发送代理通告(Agent Advertisement)消息。

(2) 接到这些消息的移动节点，知道环境中有归属代理和外区代理的存在，并确定自己是在归属网还是在外区网上。

(3) 如果移动节点收到的是归属代理发来的消息，则说明自己仍在归属网上，此时，不启动移动功能。

(4) 当移动节点检测到它移到外区网，它从外区网处获取一个临时的 IP 地址，即转交地址。

(5) 移动节点向归属代理注册，表明自己已离开归属网，把所获的关联地址通知归属代理。归属代理可以随时获取移动节点的当前位置信息。

(6) 注册完毕后，当通信端要向移动节点发送报文时，使用移动节点的固定 IP 地址。报文将被路由到移动节点的归属网，并被归属代理截获。归属代理将该报文封装，通过隧道将报文发送到移动节点所在的外区网。

由此可见，首先，移动 IP 对通信端是透明的，通信端根本不用关心当前主机是否移动，仅仅用移动主机的固定 IP 地址发送报文就可以了。

其次，移动 IP 对上层应用也是透明的。移动主机不管移动到哪里，其 IP 地址都固定不变，上层应用根本感觉不到主机位置的变化。所有变动等工作都在 IP 层完成，因此，移动 IP 是网络层的移动解决方案。

课后习题

1. IP 地址由网络号和主机号两部分组成，其中网络号表示________，主机号表示________。

2. IPv4 地址由________位二进制数组成，IPv6 地址由________位二进制数组成。

3. IP 地址 205.140.36.88 的（　　）部分表示主机号。

A. 205　　B. 205.140　　C. 88　　D. 36.88

4. IP 地址 129.66.51.37 的（　　）部分表示网络号。

A. 129.66　B. 129　　B. 129.66.51　　C. 37

5. 假设一台主机的 IP 地址为 192.168.5.121，而子网掩码为 255.255.255.248，那么该主机的网络号部分（包括子网号部分）为（　　）。

A. 192.168.5.12　　B. 192.168.5.121

C. 192.168.5.120　　D. 192.168.5.32

6. IP 地址为 140. 111. 0. 0 的 B 类网络，若要切割为 9 个子网，而且都要联上 Internet，则子网掩码要设为（　　）。

A. 255. 0. 0. 0　　B. 255. 255. 0. 0

C. 255.255.128.0　　D. 255.255.240.0

7. 255. 255. 255. 224 可能代表的是（　　）。

A. 一个 B 类网络号　　B. 一个 C 类网络中的广播

C. 一个具有子网的网络掩码　　D. 以上都不是

8. 报文交换技术说法不正确的是（　　）。

A. 报文交换采用的传送方式是“存储—转发”方式

B. 报文交换方式中数据传输的数据块其长度不限且可变

C. 报文交换可以把一个报文发送到多个目的地

D. 报文交换方式适用于语言连接或交互式终端到计算机的连接

9. 试简单说明 IP、ARP、RARP 和 ICMP 协议的作用。

10. 网络互联有何实际意义？进行网络互联时，有哪些共同的问题需要解决？

11. IP 地址分为几类？每类如何表示？IP 地址的主要特点是什么？

12. 试说明 IP 地址与硬件地址的区别。为什么要使用这两种不同的地址？

13. 回答下列问题：

（1）子网掩码为 255.255.255.0 代表什么意思？

（2）某网络现在的掩码为 255.255.255.248，问该网络能够连接多少台主机？

(3) 某一个 A 类网络和一个 B 类网络的子网号 SubnetID 分别为 16b 和 8b,问这两个网络的子网掩码有何不同?

(4) 某 A 类网络的子网掩码为 255.255.0.255,它是否为一个有效的子网掩码?

14. 试辨认以下 IP 地址的网络类别:

(1) 128.36.199.3　　(2) 21.12.240.17

(3) 183.194.76.253　　(4) 192.12.69.24

(5) 89.3.0.1　　(6) 200.3.6.2

15. IP 数据报中的首部检验和并不检验数据报中的数据,这样做的最大好处是什么?坏处是什么?

16. 在因特网上的一个 B 类地址的子网掩码是 255.255.240.0。试问在其中每一个子网上的主机数最多是多少台?

17. 某网络上连接的所有主机,都得到 Request time out 的显示输出,检查本地主机配置和 IP 地址 202.117.34.35,子网掩码为 255.255.0.0,默认网关为 202.117.34.1,请问问题可能出在哪里?

18. IPv6 使用 16B 地址空间,设每隔 1μs(微秒)就分配 100 万个地址,试计算大约要用多少年才能将 IP 地址空间全部用光?可以和宇宙的年龄(大约有 100 亿年)进行比较。

19. 一个数据报长度为 4000B(固定首部长度),现在经过一个网络传送,但此网络能够传送的最大长度为 1500B,试问应当划分为几个短些的数据报片?各数据报片的数据字段长度、片偏移字段和 MF 标志为何数值?

20. 某单位分配到一个 B 类 IP 地址,其 NetID 为 129.250.0.0。该单位有 4000 台机器,平均分布在 16 个不同的地点。如选用子网掩码为 255.255.255.0,试给每一个地点分配一个子网号码,并算出每个地点主机号码的最小值和最大值。

21. 已知某网络有一个地址是 167.199.170.82/27,问这个网络的网络掩码、网络前缀长度和网络后缀长度是多少?

22. 某单位分配到一个起始地址为 14.24.74.0/24 的地址块。该单位需要用到 3 个子网,它们的 3 个子网地址块的具体要求是:子网 N_1需要 120 个地址,子网 N_2需要 60 个地址,子网 N_3需要 10 个地址。请给出地址块的分配方案。

23. 已知路由器 R1 的路由表如图 6.41 所示。

地址掩码	目的网络地址	下一跳地址	路由器接口
/26	140.5.12.64	180.15.2.5	m2
/24	130.5.8.0	190.16.6.2	m1
/16	110.71.0.0	-----	m0
/16	180.15.0.0	-----	m2
/16	190.16.0.0	-----	m1
默认	默认	110.71.4.5	m0

图 6.41　路由器 R1 的路由表

试画出各网络和必要的路由器的连接拓扑,标注出必要的 IP 地址和接口,对不能确定的情况应当指明。

24. IP 数据报中的首部检验和并不检验数据报中的数据，这样做的最大好处是什么？缺点是什么？

25. 某单位分配到一个地址块 136.23.12.64/26。现在需要进一步划分为 4 个一样大的子网。试问：

（1）每个子网的网络前缀有多长？

（2）每一个子网中有多少个地址？

（3）每一个子网的地址块是什么？

（4）每一个子网可分配给主机使用的最小地址和最大地址是什么？

第7章 网络互联与互联设备

学习目的

随着网络技术的迅速发展和网络应用的迅速普及，网络规模迅速扩大，小型局域网已不能胜任网络应用的需要，由此，网络互联技术迅速发展起来。本章主要介绍网络互联的概念、原则、互联方式，网络互联设备包括中继器、集线器、网桥、交换机、路由器等。

学习要求

掌握：网络互联的概念和原则。

掌握：物理层互联设备。

掌握：数据链路层互联设备，特别是交换机的工作过程。

掌握：网络层互联设备路由器的工作原理。

7.1 网络互联

7.1.1 网络互联的概念

所谓网络互联就是利用网络互联设备，将两个或者两个以上具有独立自治能力的计算机网络连接起来，通过数据通信，扩大资源共享和信息交流的范围，以容纳更多的用户。20世纪90年代以来，局域网迅速发展并被广泛地应用，许多单位和部门都建立了局域网，网络的应用和信息的共享促进了网络向外延伸的需求。网络互联成了20世纪90年代计算机网络发展的标志。越来越多的人开始意识到，如果没有网络互联技术的支持，用于信息传输的计算机网络也会形成一个个“信息孤岛”。因此网络互联是计算机网络发展到一定阶段的必然结果。

在网络互联领域，类型相同(一般指网络拓扑结构或执行的协议相同)的网络称为同构网络，类型不同的网络称为异构网络，参与互联的网络一般统称为子网。网络互联应当包括同构网络互联、异构网络互联。从互联的范围看，主要体现为局域网与局域网(LAN/LAN)的互联、局域网与广域网(LAN/WAN)的互联、局域网之间经广域网的互联等。

7.1.2 网络互联原则和必须考虑的问题

为了保证网络互联可以顺利地进行，实施网络互联时通常应当遵循以下两条原则。

(1) 设计连接两个网络的互联设备时，不要轻易要求修改其中一个网络的网络结构、协议、硬件和软件。

不同的子网在诸多方面存在差异，具体表现在：寻址、信息传送、访问控制、连接方式等

几个方面。网络互联为了提供不同子网之间的网络通信,必须采取措施以屏蔽或者容纳这些差异。

(2) 不能因为要提高网络之间的传输性能而影响各个子网内部的传输功能和传输性能。

从应用的角度看,用户需要访问的资源主要还是集中在本子网内部。一般来说,网络之间的信息传输量远小于网络内部的信息传输量。但是,随着网络应用的推广,尤其是随着交换式以太网的广泛使用,局域网、局域网之间互联(主要是以太网之间互联)概念的区别已逐渐模糊。

网络互联主要应当考虑和解决以下的一些问题。

(1) 互联的层次问题。在OSI参考模型的哪一层提供网络互联是首先要考虑的问题。它涉及网络互联各个方面的问题。

(2) 寻址问题。不同子网具有不同的命名方式、地址结构,网络互联应当可以提供全网寻址的能力。

(3) 信息传送问题。网络互联可以在OSI参考模型的不同层进行,各层传送信息的格式不同。例如,物理层传送的是比特流,数据链路层传送的是数据帧,网络层传送的是数据分组等。实行网络互联时,对应不同的子网,传送的信息是不同的。例如,在网络层实现网络互联,对应不同的子网,分组的名称、长度、格式和对各种分组的处理时序会有所不同,互联的网络应当具有解决这种分组长度不兼容的能力。

(4) 访问控制问题。不同的子网采用了不同的访问控制方法(如以太网采用CSMA/CD令牌总线和令牌环采用令牌控制等),并由此而引申出各种时间的限制(如CSMA/CD中的冲突检测时间,以及各种网络协议中的传输确认的等待时间等),如何使得这些采用不同访问控制的网络可以彼此协调,共存于同一个“大”的网络中,是网络互联必须解决的又一个问题。

(5) 连接方式问题。不同的网络可能采用不同的连接方式,例如,X.25网络通常采用面向连接的信息传输,而大多数局域网又提供面向无连接服务,因此互联网提供的服务应当屏蔽这样的差异。

其他应当考虑的因素还包括不同子网的差错恢复机制对全网的影响,不同子网用户的接入限制、记账服务、通过互联设备的路由选择和网络流量控制等。

7.2 物理层互联设备

物理层位于OSI参考模型的最底层,它直接面向实际承担数据传输的物理媒体,即信道。物理层的传输单位为比特。物理层是指在物理媒体上为数据链路层提供一个原始比特流的物理连接。物理层协议规定了建立、维护及释放物理信道所需的机械的、电气的、功能性的和规程性的特性,确保比特流能在物理信道上进行传输。物理层设备是将DTE和DCE互联的设备。DTE是指数据终端设备,又称物理设备,如计算机、终端等;DCE则是数据通信设备或电路连接设备,如调制解调器等。

数据传输通常是由DTE到DCE,再由DCE到DTE的过程,DTE和DCE间的连接如图7.1所示。将DTE和DCE连接起来的装置称为物理层设备,如各种插头、插座。物理层

是OSI体系结构的最底层,它涉及的都是与物理信号(电信号、光信号等)有关的连接,常用的有Modem、中继器和集线器等。

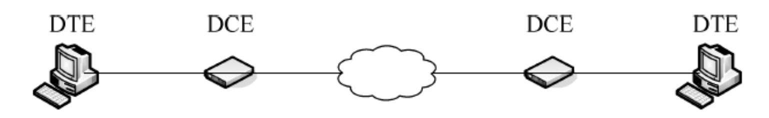


图7.1 DTE和DCE的连接

7.2.1 调制解调器

调制解调器即Modem,是调制器(Modulator)与解调器(Demodulator)的简称。Modem把数字信号转换为相应的模拟信号的过程称为“调制”。经过调制的信号传送到另一台计算机之前,接收方的Modem负责把模拟信号转换为数字信号,这个过程称为“解调”。由于目前大部分个人计算机是通过公用电话网接入计算机网络的,因而需通过调制解调器进行上述转换。

调制方式相应地有调幅、调频和调相3种。

(1)调幅:振幅调制其载波信号随着调制信号的振幅而变化。

(2)调频:载波信号的频率随着调制信号而改变。

(3)调相:相位调制有两相调制、四相调制和八相调制等方式。

1. Modem的类别

根据Modem的形态和安装方式,一般可以分为以下4类。

1)外置式Modem

外置式Modem放置于计算机的主机箱之外,如图7.2所示,通过串行口(COM口)与主机相连接。这种Modem方便灵巧、易于安装,闪烁的指示灯便于监视Modem的工作状况。但外置式Modem需要使用额外的电源与电缆。

2)内置式Modem

内置式Modem在安装时需拆开主机箱,并要对终端和COM口进行设置。这种Modem要占用主板上的扩展槽(如PCI、AMR、ACR、CNR等插槽),但无须额外的电源与电缆,如图7.3所示,且价格比外置式Modem要便宜一些。



图7.2 外置式Modem

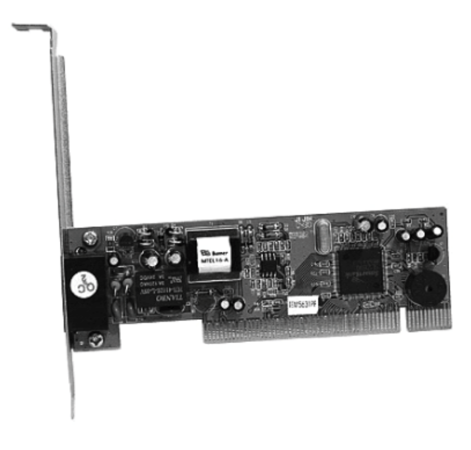


图7.3 内置式Modem

3）插卡式 Modem

插卡式 Modem 主要用于笔记本电脑，体积纤巧。配合移动电话，可方便地实现移动办公，如图 7.4 所示。

4）机架式 Modem

机架式 Modem 是将一组 Modem 集中在一个机柜中，并由统一的电源进行供电，如图 7.5 所示。机架式 Modem 主要用于企业网络、电信局、校园网、金融机构等网络的中心机房。



图 7.4 插卡式 Modem

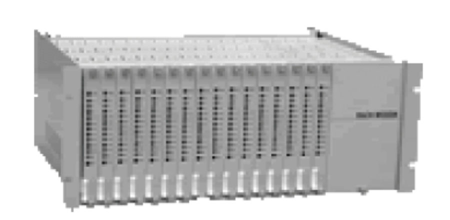


图 7.5 机架式 Modem

2. Modem 的传输模式

Modem 最初只是用于数据传输，但是随着用户需求的不断增长以及厂商之间的激烈竞争，市场上越来越多地出现了一些“二合一”“三合一”的 Modem。这些 Modem 除了可以进行数据传输以外，还具有传真和语音传输功能。

1）传真模式

通过 Modem 进行传真，除省下一台专用传真机的费用外，还有很多好处：可以直接把计算机内的文件传真到对方的计算机或传真机，而无须先把文件打印出来；可以对接收到的传真方便地进行保存或编辑；可以克服普通传真机由于使用热敏纸而造成字迹逐渐消退的问题；由于 Modem 使用了纠错的技术，传真质量比普通传真机要好，尤其是对于图形的传真更是如此。目前，传真模式(Fax Modem)大多遵循 V.29 和 V.17 传真协议，其中 V.29 支持 9600bps 的传真速率，而 V.17 则可支持 14 400bps 的传真速率。

2）语音模式

语音模式(Voice Modem)主要提供了电话录音留言和全双工免提通话功能，真正使电话与计算机融为一体。这里介绍的语音模式是 DSVD(Digital Simultaneous Voice and Data，数字式声频视频同传)。DSVD 是由 Hayes、Rockwell、U. S. Robotics、Intel 等公司在 1995 年提出的一项语音传输标准，是现有的 V.42 纠错协议的扩充。DSVD 通过采用 Digitalk 的数字式语音与数据同传技术，使 Modem 可以在普通电话线上一边进行数据传输一边进行通话。

DSVD Modem 保留了 8Kbps 的带宽(也有的 Modem 保留了 8.5Kbps 的带宽)用于语音传送，其余的带宽则用于数据传输。语音在传输前会先进行压缩，然后与需要传输的数据综合在一起，通过电话载波传送到对方用户。在接收端，Modem 先把语音与数据分离开来，

再把语音信号进行解压和数/模转换，从而实现数据/语音的同传。DSVD Modem 在远程教学、协同工作、网络游戏等方面有着广泛的应用。但由于 DSVD Modem 的价格比普通的 Voice Modem 要贵，而且要实现数据/语音同传功能时，需要对方也使用 DSVD Modem，从而在一定程度上阻碍了 DSVD Modem 的普及。

3. Modem 的传输速率

Modem 的传输速率指的是 Modem 每秒传送的数据量的大小。人们平常说的 14.4K、28.8K、33.6K、56K 等，指的就是 Modem 的传输速率，传输速率以 bps 为单位。因此，一台 33.6Kbps 的 Modem 每秒钟可以传输 33 600b 的数据。由于目前的 Modem 在传输时都对数据进行了压缩，因此 33.6K 的 Modem 的数据吞吐量理论上可以达到 115 200bps，甚至 230 400bps。

Modem 的传输速率，实际上是由 Modem 所支持的调制协议所决定的。平时在 Modem 的包装盒或说明书上看到的 V.32、V.32bis、V.34、V.34＋、V.fc 等，指的就是 Modem 所采用的调制协议。其中 V.32 是非同步/同步 4800/9600bps 全双工标准协议；V.32bis 是 V.32 的增强版，支持 14 400bps 的传输速率；V.34 是同步 28 800bps 全双工标准协议；而 V.34＋则为同步 33 600bps 全双工标准协议，以上标准都是由 ITU（国际通信联盟）所制定的；而 V.fc 则是由 Rockwell 提出的 28 800bps 调制协议，但并未得到广泛支持。

提到 Modem 的传输速率，就不能不提 56Kbps Modem。其实，56Kbps 的标准已提出多年，但由于长期以来一直存在以 Rockwell 为首的 K56Flex 和以 U.S. Robotics 为首 X2 的两种互不兼容的标准，使得 56Kbps Modem 迟迟得不到普及。在国际电信联盟的努力下，56Kbps 的标准终于统一为 ITU V9.0，众多的 Modem 生产厂商也已纷纷出台了升级措施，而真正支持 V9.0 的 Modem 早已普及。

以上所讲的传输速率，均是在理想状况下得出的，而在实际使用过程中，Modem 的传输速率往往不能达到标称值。实际的传输速率主要取决于以下几个因素。

（1）电话线路的质量。调制后的信号是由电话线进行传输的，如果电话线路质量不佳，Modem 将会降低速率以保证准确率。为此，在连接 Modem 时，要尽量减少连线长度，多余的连线要剪去，切勿绕成一圈堆放。

（2）是否有足够的带宽。如果在同一时间上网的人数很多，就会造成线路的拥挤和拥塞，Modem 的传输速率自然也会随之下降。因此，ISP 是否能提供足够的带宽非常关键。另外，避免在繁忙时段上网也是一个解决方法，尤其是在下载文件时，在繁忙时段与非繁忙时段下载所费的时间会相差几倍之多。

（3）对方的 Modem 传输速率。Modem 所支持的调制协议是向下兼容的，实际的连接速率取决于速率较低的一方。因此，如果对方的 Modem 是 14.4K 的，即使使用的是 56K 的 Modem，也只能以 14 400bps 的速率进行连接。

4. Modem 的传输协议

Modem 的传输协议包括调制协议（Modulation Protocol）、差错控制协议（Error Control Protocol）、数据压缩协议（Data Compression Protocol）和文件传输协议等。

（1）差错控制协议。随着 Modem 传输速率的不断提高，电话线路上的噪声、电流的异常突变等，都会造成数据传输出错。差错控制协议要解决的就是如何在高速传输中保证数据的准确率。差错控制协议存在着两个工业标准：MNP4 和 V4.2。其中，MNP（Microcom

Network Protocol)是 Microcom 公司制定的传输协议,包括 MNP1～MNP10。由于商业原因,Microcom 目前只公布了 MNP1～MNP5,其中 MNP4 是目前被广泛使用的差错控制协议之一。而 V4.2 则是国际电信联盟制定的 MNP4 改良版,它包含了 MNP4 和 LAP-M 两种控制算法。因此,一个使用 V4.2 协议的 Modem 可以和一个只支持 MNP4 协议的 Modem 建立无差错控制连接;而反之则不能。

(2) 数据压缩协议。为了提高数据的传输量,缩短传输时间,现在大多数 Modem 在进行传输时会先对数据进行压缩。与差错控制协议相似,数据压缩协议也存在两个工业标准:MNP5 和 V4.2bis。MNP5 采用了 Rnu-Length 编码和 Huffman 编码两种压缩算法,最大压缩比为 2∶1。而 V4.2bis 采用了 Lempel-Ziv 压缩技术,最大压缩比可达 4∶1。这就是为什么 V4.2bis 比 MNP5 要快的原因。

(3) 文件传输协议。文件传输是数据交换的主要形式。在进行文件传输时,为使文件能被正确识别和传送,需要在两台计算机之间建立统一的传输协议,这个协议包括了文件的识别、传送的起止时间、错误的判断与纠正等内容。常见的文件传输协议有以下几种。

① ASCII:这是最快的文件传输协议,但只能传送文本文件。

② Xmodem:这种古老的文件传输协议速度较慢,但由于使用了 CRC 错误检测方法,传输的准确率可高达 99.6%。

③ Ymodem:这是 Xmodem 的改良版,使用了 1024b 区段传送,速度比 Xmodem 要快。

④ Zmodem:Zmodem 采用了串流式(Streaming)传输方式,传输速率较快,而且还具有自动改变区段大小和断点续传、快速错误检测等功能,这是目前最流行的文件传输协议。

5. Modem 的安装

1) 外置式 Modem 的安装

(1) 连接电话线。把电话线的 RJ-11 插头插入 Modem 的 Line 接口,再用电话线把 Modem 上的 Phone 接口与电话机连接。

(2) 关闭计算机电源,将 Modem 所配的电缆的一端(25 针阳头端)与 Modem 连接,另一端(9 针或者 25 针插头)与主机上的 COM 口连接。

(3) 将电源变压器与 Modem 的 Power 或 AC 接口连接。接通电源后,Modem 的 MR 指示灯应常亮。如果 MR 灯不亮或不停闪烁,则表示未正确安装或 Modem 自身有故障。对于带语音功能的 Modem,还应把 Modem 的 SPK 接口与声卡上的 Line In 接口连接,当然也可直接与耳机等输出设备连接。

另外,Modem 的 MIC 接口用于连接驻极体麦克风,但最好还是把麦克风连接到声卡上。

2) 内置式 Modem 的安装

(1) 根据说明书的指示,设置好有关的跳线。由于 COM1 与 COM3、COM2 与 COM4 共用一个终端,因此通常可设置为 COM3/IRQ4 或 COM4/IRQ3。

(2) 关闭计算机电源并打开机箱,将 Modem 插入主板上任一空置的扩展槽。

(3) 连接电话线。把电话线的 RJ-11 插头插入 Modem 上的 Line 接口,再用电话线把 Modem 上的 Phone 接口与电话机连接。此时拿起电话机,应能正常拨打电话。

下面介绍一些 Modem 指示灯的含义。

(1) MR:Modem 已准备就绪,并成功通过自检。

(2) TR:终端准备就绪。

（3）SD：Modem 正在发出数据。

（4）RD：Modem 正在接收数据。

（5）OH：占机指示，Modem 正占用电话线。

（6）CD：载波检测，Modem 与对方连接成功。

（7）RI：Modem 处于自动应答状态。某些 Modem 用 AA 表示。

（8）HS：高速指示，速率大于 9600bps。

6. Modem 的芯片

Modem 的芯片就好像处理器的品牌一样有不同厂家的产品，其中占有量最大的是 Rockwell 芯片，它占全球市场份额的 70％左右，地位和处理器市场上的 Intel 差不多，目前国内大多数外置式 Modem 产品采用的是 Rockwell 芯片；其次是 TI 芯片，著名的 USR“大黑猫”就用 TI 芯片；除此以外，还有 Curiss Logic 的产品，不过使用这种芯片的外置式 Modem 比较少。总的来看，采用 Rockwell 芯片的 Modem 的性能和稳定性都比采用其他芯片的 Modem 要好。在国际市场上 TI 芯片的价格要比 Rockwell 芯片低一些。

7.2.2 中继器

由于在网络线路的传输上存在损耗，传输的数字信号或模拟信号会逐渐衰减，当衰减到一定程度时将造成信号的失真，因此会导致接收错误。中继器（Repeater）就是为解决这一问题而设计的。中继器又称重发器，工作在 OSI 体系结构的第 1 层，是物理层的设备。中继器常用于两个网络节点之间物理信号的双向转发工作，对接收信号进行再生和发送，从而增加信号传输的距离。

中继器是最简单的网络互联设备，主要完成物理层的功能，负责在两个节点的物理层上按位传递信息，完成信号的复制、调整和放大功能，以此来延长网络的长度。中继器经常用来连接同一个网络的两个或多个网段。如以太网常常利用中继器扩展总线的电缆长度，标准细缆以太网的每段长度最大是 185m，加增中继器后最多可有 5 段，因此网络最大电缆长度可以提高到 925m。

一般来说，中继器两端的网络部分是网段，而不是子网。它将两段或两段以上的以太网互联起来后，只对电缆上传输的数据信号再生放大，再重发到其他电缆段上，并不管数据中是否有错误数据或数据是不是适于该网段。对数据链路层以上的协议来说，除了中继器本身引起的一定时延以外，用中继器互联起来的若干段电缆与单根电缆并无区别。中继器的使用并非是无限的，因为网络标准中都对信号的延迟范围做了具体的规定，中继器只能在此规定范围内进行有效的工作，否则会引起网络故障。

7.2.3 集线器

集线器的英文名称为 Hub，本来是“港湾”“中心”的意思。集线器在家庭网、企业网、校园网等局域网中应用比较广泛。很多小型局域网使用带有 RJ-45 头的 5 类双绞线组成星形局域网，这种网络经常要使用到集线器。集线器的功能就是共享带宽，将局域网内各自独立的计算机连接在一起并互相通信。

1. 集线器的工作原理

工作在物理层,其实质就是一个多端口的中继器。主要功能是对接收到的信号进行再生整形放大,以扩大网络的传输距离,同时把所有节点集中在以它为中心的节点上。正因为集线器只是一个信号放大和中转的设备,所以它不具备交换功能。集线器适用于星形结构的网络布线,如果某个工作站出现问题,并不会影响整个网络的正常运行。用集线器组成的以太网是以共享总线的方式工作的,遵循"先听后说"的 CSMA/CD 协议,每台计算机在发送数据前都必须进行载波监听。只有当判定网络空闲时,才能发送数据。

如果在发送数据时检测到冲突,则该帧将被重发。当一个站点多次检测线路均为载波时,将自动放弃该帧的发送,从而造成丢包。因此,当网络中的站点数过多时,这种共享式集线器会产生大量的信号"冲突",网络的有效利用率将会大大降低。根据工程经验,采用 10Mbps 集线器的站点不宜超过 25 个,采用 100Mbps 集线器的站点不宜超过 35 个。所以,当网络较大、用户较多时,只有采用交换机才能保证每台计算机拥有较为充分的网络带宽。

2. 集线器的分类

集线器按照不同的分类标准,分为不同的种类。

(1) 按外观尺寸分类,集线器有机架式和桌面式两种。

机架式集线器是指几何尺寸符合工业规范、可以安装在 19in 标准机柜中的集线器,该类集线器以 8 口、16 口和 24 口的设备为主流。由于集线器统一放置在机柜中,既方便了集线器间的连接,也方便了对集线器的管理。有很多机架式集线器甚至没有设置电源开关,通常能够不间断地长时间工作。

桌面式集线器是指几何尺寸不符合 19in 工业规范、不能够安装在机柜中、只能直接置放于桌面的集线器。该类集线器大多遵循 8～16 口规范,也有个别 4～5 口的产品,仅适用于只有几台计算机的超小型网络,如小型办公室或家庭等。

(2) 按可提供的端口带宽分类,集线器通常有 10Mbps 集线器、100Mbps 集线器、10/100Mbps 自适应集线器及 1000Mbps 集线器 4 种。

10Mbps 集线器是指该集线器中的所有端口只能提供 10Mbps 带宽。

100Mbps 集线器是指该集线器中的所有端口只能提供 100Mbps 带宽。

10/100Mbps 自适应集线器是指该集线器可以在 10Mbps 和 100Mbps 之间进行切换。与 10/100Mbps 的自适应交换机不同,这种自适应集线器产品并不常见,由于集线器的工作原理是完全以广播的方式在所有端口上进行数据转发,所以自适应集线器的工作带宽要么是 10Mbps,要么是 100Mbps,并不能在不同的端口上以 10Mbps 或 100Mbps 的不同速率混合工作。只要有一个端口连接的是 10Mbps,则整个集线器都将工作在 10Mbps;只有连接的所有端口都是 100Mbps 时,集线器才提供 100Mbps 带宽。

1000Mbps 集线器能提供 1000Mbps 的工作带宽,但由于性价比低等原因很少能够见到。

(3) 按管理方式分类,集线器可以有亚集线器(Damp Hub)和智能集线器(Intelligent Hub)两种。

亚集线器是指不可管理的集线器,属于低端产品。

智能集线器是指能够通过 SNMP(Simple Network Management Protocol,简单网络管理协议)进行简单管理的集线,比如,启用和关闭某些端口等。这种管理大多是通过增加网

管模块来实现的。

注意，那种可以进行配置和管理的交换式集线器(Switch Hub)已经属于交换机的范畴，并非这里所说的智能集线器。

(4) 按扩展方式分类，集线器有堆叠式集线器和级联式集线器两种。

堆叠式集线器是指能够使用专门的连接线(堆叠线)，通过专用的端口(堆叠端口)将若干集线器"堆叠"在一起，从而将堆叠中的几个集线器视为一个集线器来使用和管理。级联式集线器能够通过"级联"的办法在网络中增加节点数。两个集线器进行级联时一般通过 RJ-45 交叉线进行连接；有些集线器本身提供了所谓的"级联端口"，此端口上常标有 Uplink 或 MDI 字样，通过它与其他的集线器的普通端口进行级联时只需使用通常的 RJ-45 直线就可以了。

需要特别注意的是，集线器进行级联只能是 100Mbps 的端口与 100Mbps 的端口相连，不同速率的端口不能混合连接。

3. 集线器上的端口

通常集线器主要以 RJ-45 端口连接双绞线组成以太网，所以集线器上通常会提供数量较多的 RJ-45 口。接口数通常有 8 口、12 口、16 口、24 口等几种。

早期的集线器在设计和制造上还经常考虑细缆或粗缆组网需求，所以一般集线器都同时具有 BNC 和 RJ-45 两种接口或 BNC、RJ-45 和 AUI 三种接口。

4. 集线器的缺点

集线器属于纯硬件网络底层设备，基本上不具有类似于交换机的"智能记忆"能力和"学习"能力。集线器也不具备交换机所具有的 MAC 地址表，所以它发送数据时都是没有针对性的，而是采用广播方式发送。也就是说，当它要向某节点发送数据时，不是直接把数据发送到目的节点，而是把数据包发送到与集线器相连的所有节点。

这种广播发送数据方式有如下几方面不足。

(1) 用户数据包向所有节点发送，很可能带来数据通信的不安全因素，一些别有用心的人很容易非法截获他人的数据包。

(2) 由于所有数据包都是向所有节点同时发送，加上以上所介绍的共享带宽方式，就更加可能造成网络拥塞现象，更加降低了网络执行效率。

(3) 非双工传输，网络通信效率低。集线器的同一时刻每一个端口只能进行一个方向的数据通信，而不能像交换机那样进行双向双工传输，网络执行效率低，不能满足较大型网络通信需求。

集线器价格非常便宜、组网灵活，所以以往经常使用。但随着交换机价格的不断下降，仅有的价格优势已不再明显，集线器的市场越来越小，基本上已经被淘汰，所以目前在市面上已很少能买到集线器这种产品了。

7.3 数据链路层互联设备

数据链路层是 OSI 体系结构中构建局域网的非常重要的层次之一。数据链路层的互联设备主要有网卡和交换机等。

7.3.1　网卡

网络适配器(Network Interface Adapter,NIA)又称网络接口卡(Network Interface Card,NIC),简称网卡。网卡实现联网计算机和网络电缆之间的物理连接,为计算机之间相互通信提供一条物理通道,并通过这条通道进行高速数据传输。

由于网络技术的迅速发展和应用普及,网卡也从原先的独立板卡越来越多地集成到了计算机的主板上。

1. 网卡的功能

网卡是计算机网络中最基本的元素。在局域网中,每一台联网计算机都需要安装一块或多块网卡,通过介质连接器将计算机接入网络电缆系统。如果有一台计算机没有网卡,那么这台计算机将不能和其他计算机通信,也就是说,这台计算机和网络是孤立的。

网卡完成物理层和数据链路层的大部分功能,包括网卡与网络电缆的物理连接、介质访问控制(如 CSMA/CD)、帧的装拆、帧的发送与接收、错误校验、数据信号的编/解码(如曼彻斯特编码)、数据收发缓冲、数据的串并行转换等功能。

网卡必须具备两大技术:网卡驱动程序和 I/O 技术。驱动程序使网卡与网络操作系统兼容,实现 PC 与网络的通信。I/O 技术可以通过数据总线实现 PC 和网卡之间的通信。

2. 网卡的分类

(1) 根据网络技术的不同分类。包括大家所熟知的以太网网卡、ATM 网卡、令牌环网卡等,不同的网卡使用完全不同的介质技术。目前,由于以太网占据了局域网技术的绝对主流,所以本书主要以以太网卡为据。

(2) 根据接口类型和传输介质的不同,以太网卡分为 AUI 接口(粗缆接口)、BNC 接口(细缆接口)和 RJ-45 接口(双绞线接口)3 种接口类型。目前市场上主要是 RJ-45 型,前两种已基本被淘汰。

(3) 根据网卡的总线类型,主要分为 ISA 网卡、EISA 网卡和 PCI 网卡 3 类。其中,前两种也基本被淘汰,而以 PCI 网卡最为常见。

(4) 根据网卡的速度分,通常有 10Mbps、100Mbps、10/100Mbps 及千兆网卡。10Mbps、100Mbps 目前都不太容易见到了,市场的主流是 10/100Mbps 自适应网卡,即网卡可以与远端网络设备(集线器或交换机)自动协商,确定当前的可用速率是 10Mbps 还是 100Mbps。就整体价格和技术发展而言,千兆以太网已经有所应用但普及尚需时日,而 10Mbps 的时代已经远去了。千兆网卡目前主要应用在服务器中。

(5) 根据工作对象的不同,网卡还可以分为普通工作站网卡和服务器专用网卡。服务器专用网卡是根据服务器的工作特点而专门设计的,对内存和 CPU 的占用率很低、性能很好,但价格较高。

另外,有的网卡在 BootROM 上做文章,加入了防病毒功能;有的网卡则与主板配合,借助一定的软件,实现 Wake On LAN(远程唤醒)功能,可以通过网络远程启动计算机;在笔记本电脑中,有时会用到 PCMCI A 接口类型的网卡,甚至还有 USB 接口的网卡;在一些服务器上,有时还会看到用于光纤连接的光纤网卡等。

7.3.2 网桥

网桥(Bridge)也叫桥接器,工作在数据链路层。网桥有在不同网段之间再生信号的功能,它可以有效地连接两个局域网,根据 MAC 地址来转发帧,使本地通信限制在本网段内,并转发相应的信号至另一网段,网桥通常用于连接数量不多的、同一类型的网段。可以看作一个"低层的路由器"(路由器工作在网络层,根据网络地址如 IP 地址进行转发)。

网桥的功能在延长网络跨度上类似于中继器,然而它能提供智能化的连接服务,即根据帧的目的地址处于哪一网段来进行转发和滤除。网桥对站点所处网段的了解是靠"自学习"实现的。使用网桥进行互联克服了物理限制,这意味着构成局域网的数据站总数和网段数很容易扩充。

网桥纳入存储和转发功能可使其适应于连接使用不同 MAC 协议的两个 LAN,因而构成一个不同 LAN 混联在一起的混合网络环境。网桥的中继功能仅仅依赖于 MAC 帧的地址,因而对高层协议完全透明。网桥将一个较大的 LAN 分成段,有利于改善可靠性、可用性和安全性。

7.3.3 交换机

集线器作为第一种广泛应用的网络设备,在各大局域网中应用非常广泛。但随着网络传输介质类型的日益丰富,图形、图像及各种流媒体等多媒体内容的出现,人们对提高网络传输速率和传输性能的需求日益增长。集线器由于共享介质传输、单工数据操作和广播数据发送方式等固有特性都决定了难满足用户在传输速率和传输性能上的要求。在用户的需求推动下、在全球各大网络设备开发商的努力下,一种更新、更实用的设备——交换机出现了,如图 7.6 所示。



图 7.6 交换机

1. 交换机的产生与发展

交换机的英文名称为 Switch。它是集线器的升级换代产品,从外观上来看,它与集线器基本上没有多大区别,都是带有多个端口的长方形盒状体。交换机按照通信两端传输信息的需要,用人工指定或设备自动完成的方法把要传输的信息送到符合要求的相应链路上。广义的交换机就是一种在通信系统中完成信息交换功能的设备。

"交换"和"交换机"最早起源于电话通信系统(PSTN)。早期电话由接线生完成线路的转接以实现电话两端的通信,接线生实际在扮演着"交换机"的角色,只不过它是一种人工的电话交换系统,不是自动的,也不是我们今天要谈的计算机网络中用的交换机,但我们现在所说的网络交换机就是在电话交换机技术上发展而来的。

在计算机网络系统中,交换概念的提出是对共享工作模式的重大改进。我们知道集线器(Hub)是一种共享介质的网络设备,而且 Hub 本身不能识别目的地址,是采用广播方式向所有节点发送。即当同一局域网内的 A 主机向 B 主机传输数据时,数据包在以 Hub 为架构的网络上是以广播方式传输的,对网络上所有节点同时发送同一信息,然后再由每一台

终端通过验证数据报报头中的地址信息来确定是否接收。在这种方式下我们知道很容易造成网络拥塞,因为其实接收数据的一般来说只有一个终端节点,而现在对所有节点都发送,那么绝大部分数据流量是无效的,这样就造成整个网络数据传输速率相当低。

另一方面由于每个节点都能监听到所发送的数据包,显然容易出现一些不安全因素。交换机完全克服了集线器的上述种种不足,所以在短时间内得到业界广泛的认可和应用。交换机技术得到了飞速发展,数据传输速率快速增加。目前千兆级的交换机在企业骨干网络中早已得到广泛应用,最快的以太网交换机端口带宽可达到10Gbps(万兆)。

2. 交换机的工作原理

交换机拥有一条很高带宽的背板总线和内部交换矩阵。交换机的所有端口都挂接在这条背板总线上。控制电路收到数据帧以后,处理端口会查找内存中的MAC地址(网卡的硬件地址)对照表(这张表格是集线器产品所没有的)以确定目的MAC的网卡挂接在哪个端口上,然后通过内部交换矩阵直接将数据帧迅速传送到目的节点,而不是所有节点,目的MAC若不存在才广播到所有的端口。这种方式就是"交换"。可以明显地看出"交换"一方面效率比较高,不会浪费网络资源,只是对目的地址发送数据,不易产生网络拥塞;另一方面数据传输也比较安全,因为它不是对所有节点都同时发送,其他节点很难监听到所发送的信息。这也是交换机为什么会很快取代集线器的重要原因之一。

使用交换机也可以把网络"分段",通过地址对照表,交换机只允许必要的网络流量通过。通过交换机的过滤和转发,可以有效地隔离广播风暴,减少错误包的出现,避免共享冲突。

交换机在同一时刻还可以进行多个端口对之间的数据传输。每一端口都相对独立,连接在其上的网络设备独自享有全部的带宽,无须与其他设备竞争使用。当节点A向节点D发送数据时,节点B可同时向节点C发送数据,而且这两个传输都享有网络的全部带宽,都有着自己的虚拟连接。假设使用的是100Mbps以太网的交换机,那么该交换机的总流量就等于2×100Mbps=200Mbps,而使用100Mbps共享式Hub时,总流量不会超出100Mbps。

总之,交换机是一种基于MAC地址识别,能完成封装转发数据包功能的数据链路层设备。交换机可以"学习"MAC地址,并存放在内部地址表中,通过在数据帧的始发者和目标接收者之间建立临时的交换路径,使数据帧直接由源地址到达目的地址。

3. 交换机的功能

交换机的主要功能包括物理编址、网络拓扑结构、错误校验、帧序列以及流量控制。目前交换机还具备了一些新的功能,如对VLAN(虚拟局域网)的支持、对链路汇聚的支持等。交换机除了能够连接同种类型的网络之外,还可以在不同类型的网络(如以太网和快速以太网)之间起到互联作用。如今许多交换机都能够提供支持百兆的快速以太网和千兆以太网,甚至万兆以太网等的高速连接端口,有些还有ATMF、DDI等其他类型的端口,用于连接网络中的其他交换机或者为带宽占用量大的关键服务器提供附加带宽。

一般来说,交换机的每个端口都用来连接一个独立的网段,但由于交换机价格迅速下降,为了提供更快的接入速率和更好的网络性能,可以把一些普通计算机直接连接到交换机的端口上,而重要的计算机等设备可以连接到更高速的交换端口上。这样,网络上的普通计算机、关键服务器和重要用户都能拥有更快的接入速率,支持更高的数据流量。

4. 交换机的3种交换方式

(1) 直通式(Cut Though)。直通式以太网交换机可以理解为在各端口间是纵横交叉的

线路矩阵电话交换机。它在输入端口检测到一个数据帧时,检查该帧帧头,获取帧的目的地址,启动内部的动态查找表转换成相应的输出端口,在输入与输出交叉处接通,把数据帧直通到相应的端口,实现交换功能。由于不需要存储,因此延迟非常小、交换非常快,这是它的优点。缺点是,由于数据帧内容并没有被以太网交换机保存下来,所以无法检查所传送的数据帧是否有误,不能提供错误检测能力。由于没有缓存,不能将具有不同速率的输入/输出端口直接接通,而且容易丢帧。

(2) 存储转发(Store&Forward)。存储转发方式是计算机网络领域应用最为广泛的方式。它把输入端口的数据帧先存储起来,然后进行 CRC 校验,在对错误帧处理后才取出数据帧的目的地址,通过查找表转换成输出端口,从而送出数据包。正因为如此,存储转发方式在数据处理时延时大,这是它的不足,但是它可以对进入交换机的数据帧进行错误检测,有效地改善网络性能。尤其重要的是,它可以支持不同速度的端口间的转换,保持高速端口与低速端口间的协调工作。

(3) 碎片隔离(Fragment Free)。这是介于前两者之间的一种解决方案。它检查数据帧的长度是否达到最小要求的 64B,如果小于 64B,说明是错误帧,则丢弃该帧;如果大于 64B,则发送该帧。这种方式也不提供数据校验,它的数据处理速度比存储转发方式快,但比直通式慢。

5. 交换机的分类

(1) 从广义上来看,交换机分为两种:广域网交换机和局域网交换机。广域网交换机主要应用于电信领域,提供通信用的基础平台;而局域网交换机则应用于局域网络,用于连接终端设备,如 PC、网络打印机等。

(2) 从传输介质和传输速率上可分为以太网交换机、快速以太网交换机、千兆以太网交换机、万兆以太网交换机、FDDI 交换机、ATM 交换机和令牌环交换机等。

(3) 从应用规模上可分为企业级交换机、部门级交换机、工作组交换机、桌面型交换机等。各厂商划分的尺度并不完全一致。一般来讲,企业级交换机都是机架式的、可扩展的,部门级交换机可以是机架式(插槽数较少),也可以是固定配置式,而工作组交换机为固定配置式(功能较为简单)。另一方面,从应用的规模来看,作为骨干交换机时,支持 500 个信息点以上大型企业应用的交换机为企业级交换机,支持 300 个信息点以下中型企业的交换机为部门级交换机,而支持 100 个信息点以内的交换机为工作组交换机。桌面型交换机是最常见的一种低档交换机,它区别于其他交换机的一个主要特点是支持的每端口 MAC 地址很少,通常端口数为 5~8 个,一般不超过 12 个,只具备最基本的交换机特性,当然价格也很便宜,比较适合 SOHO 环境。

(4) 根据交换机端口结构划分,可以分成固定端口交换机和模块化交换机,以及两者兼顾的在提供基本固定端口的基础上再配备一定的扩展插槽或模块。固定端口交换机所带有的端口是固定的,以 24 口最为常见,此外还有 8 口、16 口、48 口等多种。模块化交换机在价格上要贵很多,但拥有更强的功能和性能、更大的灵活性和可扩充性等,可以根据需要选择不同数量、不同速率、不同接口类型的模块。而且,模块化交换机大多有很强的容错能力,支持交换模块的冗余备份,往往还拥有可热插拔的双电源等。在选择交换机时,应根据需要和经费等情况考虑。

(5) 根据交换机是否支持网管功能划分,可以分成网管型交换机和非网管型交换机。

网管型交换机的任务就是使所有的网络资源处于良好的状态。网管型交换机产品提供了基于终端控制口(Console)、基于Web以及支持Telnet远程登录等多种网络管理方式,网络管理员可以方便地通过本地的或远程的方式实时监控交换机的工作状态、网络运行状况等,对交换机所有端口的工作状态和工作模式(如端口速率、双工/半双工、Access/Trunk关闭/启用等)实施管理。有的网管型交换机还提供QoS(Quality of Service)访问控制等多功能。

6. 地址学习

在以太网中,计算机发送的任何数据被封装成为以太网数据帧时在帧头加入源和目的的MAC地址信息。交换机就是根据这个信息来判断其各端口所连接的设备的,其实质是保存一份供交换机随时查询设备所在端口的"地址表",即我们说的"端口地址表"。简单地说,交换机可以记住在一个接口上所收到的数据帧的源MAC地址,并将此MAC地址与接收端口的对应关系存储到MAC地址表中。

交换机采用的算法是逆向学习法(Backward Learning)。查看源地址即可知道在哪个网段上可访问哪台机器,于是在MAC地址表中添上一项。地址学习的过程如图7.7所示。

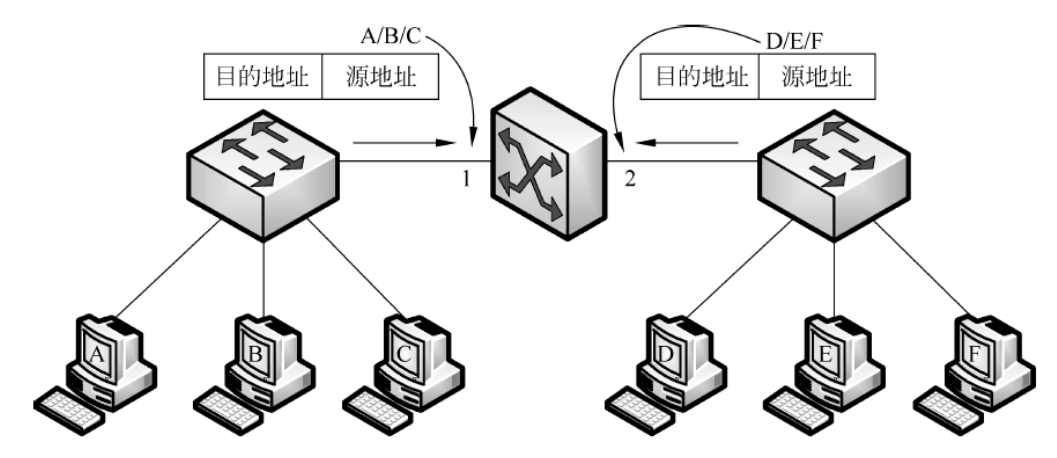


图7.7　地址学习的过程

图7.7所示的环境中,交换机根据来自1和2端口的数据帧源地址,在稳定之后将可以形成如表7.1所示的MAC地址表。

表7.1　MAC地址表

地址	端口
A/B/C	1
D/E/F	2

在交换机加电启动之初,MAC地址表为空。由于交换机不知道任何目的地的位置,因而采用扩散算法(Flooding Algorithm):把每个到来的目的地不明的帧输出到此交换机的所有其他端口并通过这些端口发送到其所连接的每一个物理网段中(除了发送该帧的物理网段外)。随着发送数据帧的站点的逐渐增多,一段时间之后,交换机将了解每个站点与交换机端口的对应关系找到相应的端口进行定向的发送。

当计算机和交换机重启、断电或迁移时,网络的拓扑结构会随之改变。为了处理动态拓扑问题,每当增加MAC地址表项时,均在该项中注明帧的到达时间。每当目的地已在表中的帧到达时,将以当前时间更新该项,如果较长时间不更新,交换机就认为该条项目过期而

删除它。这样,从物理网段上取下一台计算机,并在别处重新连接到网段上,在几分钟内,它即可重新开始正常工作而无须人工干涉。这个算法同时也意味着,如果机器在一段时间内无动作,那么发给它的帧不得不被发送到各个端口,一直到它自己发送出数据帧为止。

例:交换机学习功能示例(网络拓扑如图 7.8 所示)。

计算机 A、B 和 C 通过一台共享式集线器相连,其中集线器的一个端口与交换机的 E1 端口相连,交换机的 E2 端口连接一台网络服务器 F。

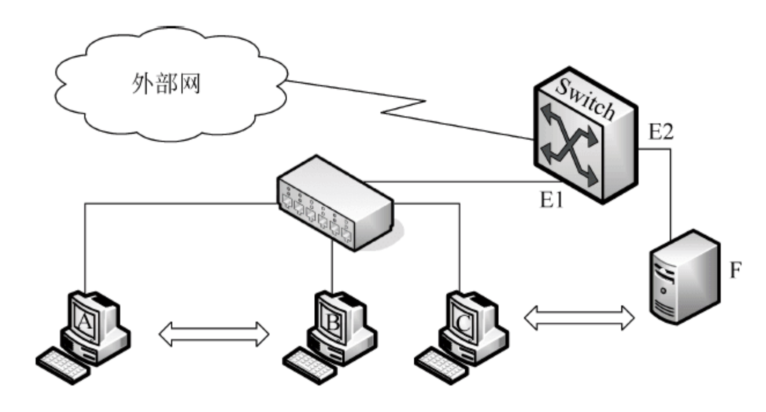


图 7.8 交换机的地址学习功能

当交换机加电自检并成功后,交换机即开始监测各个端口下连接的设备。当 A 第一次向同网段的 B 节点发送一次单播数据时,由于 A、B 及 C 都处于同一个共享网络段,因此这个数据可以被 B 和 C 收到。当然交换机的 E1 端口也可以收到这个数据,此帧的目的地址是 B 的 MAC 地址,而源地址是 A 的 MAC 地址,交换机在查看其"MAC 端口地址表"后发现没有对应表项,首先将帧的源 MAC 地址 A 与 E1 端口对应起来,然后将此数据"扩散"到除 E1 之外的所有端口,包括途中的 E2。根据分析交换机在首次转发一个单播数据时,尽管这个数据是属于某个端口内部的,其他端口的终端也可以收到这个数据,此时交换机的工作方式就如同一个集线器。

如果 B 接收到数据后回应给 A 一个消息,这个消息也同样会经过 A、B、C 的共享链路到达 A、C,也同时到达 E1 接口。但此时由于交换机在查看其"MAC 端口地址表"时,发现了帧的目的地址在表中是有对应表项的,因此不会像第一个数据一样向所有端口扩散,而是根据查询的结果进行判断再决定是过滤还是转发数据。

C 给服务器 F 发送一个服务请求时,由于共享链路的存在,A、B 都可以收到这个数据帧的同时交换机的 E1 端口也可以收到,当交换机发现其现有的"MAC 端口地址表"没有帧目的地址的对应表项也没有与源地址对应的表项后,交换机首先将其"MAC 端口地址表"添加有关源地址与端口的对应关系,然后将数据以"扩散"的方式发送到除 E1 端口之外的所有端口。这样 F 服务器一定会收到。

当 F 回应数据给 C 时,交换机通过网络介质收到数据帧,查看"MAC 端口地址表"以决定如何处理数据帧,当交换机发现在其表中不存在源 F 对应的表项时,它首先将 F 与端口 E2 对应,然后再根据目的 C 对应的端口 E1 转发数据。值得注意的是,当数据从 E1 端口发往 C 时,由于 A、B 和 C 是通过共享设备(集线器)相连,它们都将收到这个数据,与 C 不同的是,A 与 B 将不会处理这样的数据。

此时交换机对数据已经完成了全部地址学习。经过上面的数据发送和接收过程,交换

机的 MAC 地址表已经有了如表 7.2 所示的 4 条表项。

表 7.2　交换机 MAC 地址表内容示例

设备	端口	MAC
A	E1	01-11-5A-00-43-7E
B	E1	01-11-5A-00-74-A0
C	E1	01-11-51-00-E0-4F
F	E2	01-11-51-00-3C-C5

因此综上所述，当帧到达端口时，帧的出口选择过程取决于源所在的端口（源端口）和目的地所在的端口（目的端口）是否相同，总结起来有以下 3 点。

（1）源端口和目的端口相同，则丢弃该帧，即过滤；

（2）源端口和目的端口不同，则转发该帧，即转发；

（3）目的端口未知，则进行广播。

当交换机某个接口上收到数据，就会查看目的 MAC，并检查 MAC 地址表，对于交换机认为源和目的在同一个端口的数据帧，它将认为不应该发送到其他端口，影响其他端口的网络数据传输。对于这样的数据，交换机将过滤（即丢弃），以避免本地数据帧影响网络上的正常通信。过滤决定如图 7.9 所示。

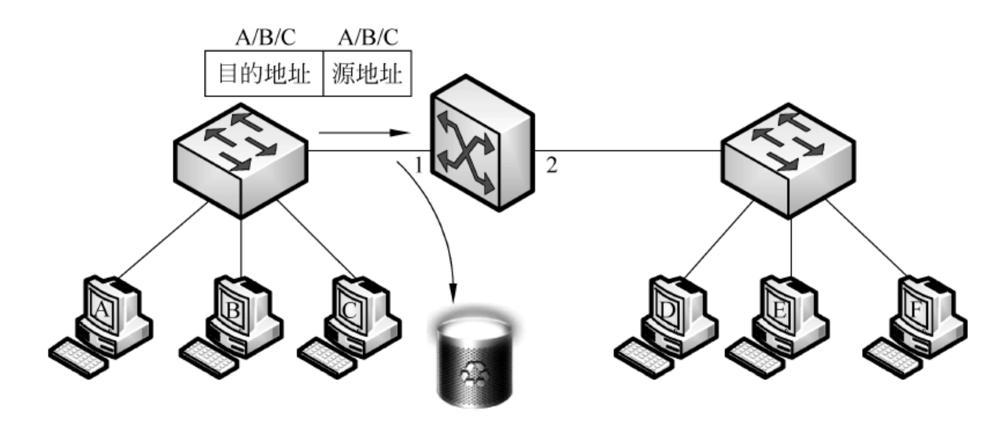


图 7.9　交换机的数据过滤

当交换机接收到一个数据帧，它的目的地址对应端口与接收端口不同，此时交换机认为有必要将数据进行转发，这就是交换机的转发过程。由于交换机仅将数据帧发送给目的地址，而不是发送给网段内的所有地址，可以有效地减少网段内的拥塞。转发决定如图 7.10 所示。

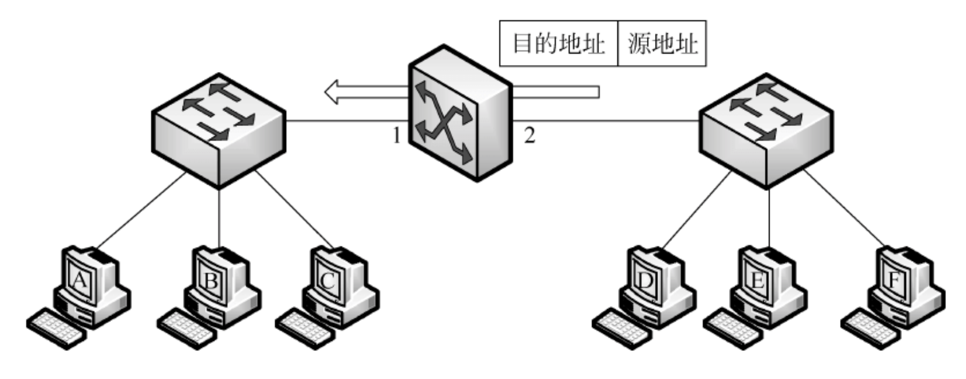


图 7.10　交换机的数据转发

7.4 网络层互联设备

7.4.1 路由器

路由器(Router)是因特网上非常重要的设备之一,正是遍布世界各地的不计其数的路由器构成了因特网这个在我们身边日夜不停地运转的巨型信息网络的“桥梁”。如图 7.11 所示,路由器与交换机拥有相似的外观。



图 7.11 路由器

1. 路由器的概念

Internet 核心通信机制是一种“存储转发”的数据传输模型。在这种通信机制下,网络上流动的所有数据都是以数据包(Packct)形式发送、传输和接收处理。接入因特网的任何一台计算机要与别的机器相互通信并交换信息就必须拥有一个唯一的网络“地址”。数据并不是从“出发点”直接就被传送到“目的地”的,而是在传送之前按照特定的标准划分成长度一定的片段——数据包。每一个数据包中都加入了目的计算机的网络地址,这就好比套上了一个写好收件人地址的信封,这样的数据包在网络上传输时才不会“迷路”。这些数据包在到达目的地之前必须经过因特网上为数众多的通信设备或者计算机的层层转发、接力传递。古代驿站的运作情形就是这个过程的一个形象比喻,在因特网上,路由器正是扮演着转发数据包“驿站”的角色。

多年来,路由器的发展有起有伏。20 世纪 90 年代中期,传统路由器成为制约因特网发展的瓶颈。ATM 交换机取而代之,成为 IP 骨干网的核心,路由器变成了配角。进入 90 年代末期,Internet 规模进一步扩大,流量每半年翻一番,ATM 又成为瓶颈,路由器东山再起。千兆级路由交换机在 1997 年面世后,人们又开始以千兆路由交换机取代 ATM 交换机,架构以路由器为核心的骨干网。

路由器连接多个网络或网段,是典型的网络层设备,在 OSI/RM 被称为中介系统,完成网络层中继或第 3 层中继的任务。

2. 路由器的处理机制

路由器连接多个逻辑上分开的网络,所谓逻辑网络,是代表一个单独的网络或者一个子网。当数据从一个子网传输到另一个子网时,可通过路由器来完成。因此,路由器具有判断网络地址和选择路径的功能,它能在多网络互联环境中建立灵活的连接,可用完全不同的数据分组和介质访问方法连接各种子网。路由器只接收源站或其他路由器的信息,属网络层的一种互联设备,它不关心各子网使用的硬件设备,只要求运行与网络层协议相一致的软件。这种方式有效地解决了异构网络的互联问题。

路由器的主要工作就是为经过路由器的每个数据包寻找一条最佳传输路径,并将该数据有效地传送到目的站点。由此可见,选择最佳路径的策略即路由选择算法是路由器的关键所在。为了完成这项工作,在路由器中保存着各种传输路径的相关数据——路由表(Routing Table)路由。路由表中保存着子网的标识信息、网上路由器的个数和下一台路由

器的名称等内容。路由表可以是静态的、由系统管理员固定设置好的,也可以由系统动态修改,即由路由器自动调整,或者由主机控制。

(1) 静态路由表。由系统管理员事先设置好的固定路由表。一般是在系统安装时就根据网络的配置情况预先设定的。

(2) 动态路由表。是路由器根据网络系统的运行情况自动调整的路径表。路由器根据路由选择协议(Routing Protocol)提供的、自动学习和记忆网络等运行情况,自动计算数据传输的最佳路径。

3. 路由器的功能

路由器有两大典型功能,即数据通道功能和控制功能。数据通道功能包括转发决定(路由选择)、背板转发以及输出链路调度等,一般由特定的硬件来完成。控制功能一般用软件来实现,包括与相邻路由器之间的信息交换、系统配置、系统管理等。

路由器还具有网络流量控制功能。有的路由器仅支持单一协议,但大部分路由器可以支持多种协议的传输,即多协议路由器。由于每一种协议都有自己的规则,要在一台路由器中完成多种协议的算法,势必会降低路由器的性能。因此,支持多协议路由器通常性能相对较低。用户购买路由器时需要根据自己的实际情况选择自己需要网络协议的路由器。近年来,出现的交换路由器(三层交换机)产品,从本质上来说它不是什么新技术,而是为了提高通信能力,把交换机的原理组合到路由器中,使数据传输更快、更好。

路由器的功能具体来说表现在以下几个方面。

(1) 在网络间截获发送到远地网段的报文,起转发的作用。

(2) 选择最合理的路由,引导通信。为了实现这一功能,路由器要按照某种路由通信协议查找路由表。路由表中列出整个互联网中包含的各个节点,以及节点间的路径情况和相关传输成本。如果到特定的节点有一条以上的路径,则基于预先确定的准则选择最优(最经济)的路径。由于各种网络段和其相互连接的情况可能发生变化,因此路由情况的信息需要及时更新,这是由所使用的路由信息协议规定的定时更新或者按变化情况更新来完成。网络中的每台路由器按照这一规则动态地更新它所保持的路由表,以便保持有效的路由信息。

(3) 路由器在转发报文的过程中,为了便于在网络间传送报文,按照预定的规则把大的数据包分解成适当大小的数据包,到达目的地后再把分解的数据包重新组装成原有的形式。

(4) 多协议路由器可以连接使用不同通信协议的网络段。

(5) 路由器的主要任务是把通信引导到目的网络,然后到达特定的节点地址。

4. 路由器的分类

路由器产品按照不同的划分标准有多种类型。常见的分类主要如下。

(1) 按性能档次分为高、中、低端路由器。通常将背板吞吐量大于40Gbps的路由器称为高端路由器,25Gbps～40Gbps的称为中端路由器,低于25Gbps的称为低端路由器。当然这只是一种宏观上的划分标准,各厂家划分并不一致,实际上路由器档次的划分不仅以吞吐量为依据,而是有一个综合指标的。

衡量路由器能力的还有一个重要指标——包转发率,也称端口吞吐量,是指路由器在某端口进行的数据包转发能力,单位通常使用p/s(包每秒)来衡量。一般来讲,低端路由器包转发率只有几K到几十Kp/s,而高端路由器则能达到几十Mp/s(百万包每秒)甚至上百

Mp/s。如果小型办公网使用,则选购转发速率较低的低端路由器即可;如果是大中型企业部门应用,就要严格注重这个指标,建议性能越高越好。

(2) 从结构上分为“模块化路由器”和“非模块化路由器”。模块化结构可以灵活地配置路由器,以适应企业不断增加的业务需求,非模块化的就只能提供固定的端口。通常中高端路由器为模块化结构,低端路由器为非模块化结构。

(3) 从功能上分为“骨干级路由器”“企业级路由器”和“接入级路由器”。骨干级路由器是实现企业级网络互联的关键设备,它数据吞吐量较大,非常重要。对骨干级路由器的基本性能要求是高速率和高可靠性。为了获得高可靠性,网络系统普遍采用诸如热备份、双电源、双数据通路等传统冗余技术。

企业级路由器连接许多终端系统,但系统相对简单,且数据流量较小,对这类路由器的要求是以尽量便宜的方法实现尽可能多的端点互联,同时还要求能够支持不同的服务质量。

接入级路由器主要应用于连接家庭或 ISP 内的小型企业客户群体。

(4) 按所处网络位置划分,通常把路由器划分为边界路由器和中间节点路由器。很明显边界路由器是处于网络边缘,用于不同网络路由器的连接;而中间节点路由器则处于网络的中间,通常用于连接不同网络,起到一个数据转发的桥梁作用。由于各自所处的网络位置有所不同,其主要性能也就有相应的侧重,如中间节点路由器要面对各种各样的网络。如何识别这些网络中的各节点呢?靠的就是这些中间节点路由器的 MAC 地址记忆功能。基于上述原因,选择中间节点路由器时就需要在 MAC 地址记忆功能更加注重,也就是要求选择缓存更大、MAC 地址记忆能力较强的路由器。但是边界路由器由于可能要同时接收来自许多不同网络路由器发来的数据,所以就要求这种边界路由器的背板带宽要足够宽,当然这也要由边界路由器所处的网络环境而定。

(5) 从性能上可分为“线速路由器”和“非线速路由器”。所谓“线速路由器”,就是完全可以按传输介质带宽进行通畅传输,基本上没有间断和延时。通常线速路由器是高端路由器,具有非常高的端口带宽和数据转发能力,能以介质速率转发数据包;中低端路由器是非线速路由器。但是一些新的宽带接入路由器也有线速转发能力。

7.4.2 三层交换机

我们经常把三层交换机理解为两层交换机+路由器,也就是带有路由功能的交换机,从概念、原理的意义上讲并无不妥,但在工程上还是有比较大的差别。从外观上看,三层交换机与两层交换机差不多,但因三层交换机大多用于网络的主干,功能更多更强大,因此往往会大一些,如图 7.12 所示。



图 7.12 三层交换机

1. 三层交换机的作用

三层交换机的特点主要在于它在具备一定路由能力的同时还具有强大的两层数据转发能力,它在处理一组数据的转发时只检查第一个数据包以选择必要的路由,其余的数据就全部交由两层交换

去处理，这使其在网络主干和子网的连接上能够发挥出最大的作用。

1）应用于网络主干

在校园网、大型企业网、城域网中，在网络的主干和汇聚层都要三层交换机作为核心实施数据的转发，否则整个网络数百、数千甚至更多的计算机都在一个子网中，不仅毫无安全性可言，也会因为无法分割广播域而无法隔离广播风暴。

如果采用传统的路由器，虽然可以隔离广播域，但是性能得不到保障。而三层交换机的性能非常高，既有三层路由的功能，又具有两层交换的速度。两层交换基于 MAC 寻址，三层交换则转发基于第 3 层地址的业务流；除了必要的路由决定过程外，大部分数据转发过程由两层交换处理，大大提高了数据包转发的效率。

三层交换机通过使用硬件交换机制实现了 IP 的路由功能，其优化的路由软件使得路由过程效率提高，解决了传统路由器软件路由的速度问题。因此可以说，三层交换机具有“路由器的功能、交换机的性能”。

2）应用于子网连接

同一网络上的计算机如果超过一定数量(通常在 200 台左右，视通信协议而定)，就很可能会因为网络上大量的广播包而导致网络传输速率低下。为了避免在大型交换机上进行广播所引起的广播风暴，可将其进一步划分为多个虚拟网(VLAN)。这样做导致的一个尖锐的问题是 VLAN 间的通信必须通过路由器来实现。但是传统路由器难以胜任 VLAN 之间的通信任务，因为相对于局域网的网络流量来说，传统的普通路由器的数据转发能力太弱，千兆级路由器性能虽然大大提升，但价格又难以接受。

如果使用三层交换机上的千兆端口或百兆端口连接不同的子网或 VLAN，就能在保证性能的前提下，经济地解决子网划分后子网之间必须依赖路由器进行通信的问题。因此三层交换机是连接子网的理想设备。

2. 三层交换机与路由器的主要区别

之所以有人搞不清三层交换机和路由器之间的区别，最根本的原因就是三层交换机也具有“路由”功能，与传统路由器的路由功能总体上是一致的。虽然如此，三层交换机与路由器还是存在着相当大的本质区别。

1）主要功能不同

虽然三层交换机与路由器都具有路由功能，但不能因此把它们等同起来。三层交换机仍是“交换机”，只不过它兼具了一些基本的路由功能，主要功能仍是数据交换。而路由器仅具有路由转发这一种主要功能。

2）主要的适用环境不一样

三层交换机所面对的应用主要是简单的局域网连接，因此三层交换机的路由功能通常比较简单，远没有路由器那么复杂。它在局域网中的用途主要是提供快速数据交换功能，满足局域网数据交换频繁的特点。而路由器则不同，它的设计初衷就是为了连接不同类型的网络，虽然也适用于局域网之间的连接，但它的路由功能更多地体现在不同类型网络之间的互联上，如局域网与广域网之间的连接、不同协议的网络之间的连接等。它最主要的功能就是路由转发，解决好各种复杂路由路径网络的连接就是它的最终目的，优势在于选择最佳路由、负荷分担、链路备份及与其他网络进行路由信息的交换。为了连接各种类型的网络，路由器的接口类型非常丰富，而三层交换机则一般仅具有同类型的局域网接口，比较简单。

3）性能体现不一样

从技术上讲，路由器和三层交换机在数据包交换操作上存在着明显区别。路由器一般由基于微处理器的软件路由引擎执行数据包交换，而三层交换机通过硬件执行数据包交换。三层交换机在对第一个数据流进行路由后，它将会产生一个 MAC 地址与 IP 地址的映射表，当同样的数据流再次通过时，将根据此表直接从两层通过而不是再次路由，从而消除了路由器进行路由选择造成的延迟，提高了数据包转发的效率；同时，三层交换机的路由查找是针对数据流的，它利用缓存技术很容易通过 ASIC 技术来实现，既可以实现快速转发，又能够大大节约制造成本。而路由器的转发采用最长匹配的方式，算法复杂，难以通过硬件直接实现，而通常使用软件完成，转发速率较低。

我们可以把路由器比作田径场上的全能运动员，能跑能跳也能投，但跑不很快（各种协议都支持，但转发能力一般）；而三层交换机就像百米飞人，能跑，接力跑也行，其他项目就不怎么样了（支持的协议虽然少，但转发能力特别强）。所以，三层交换机与路由器之间还是存在着非常大的本质区别。从整体性能上比较，三层交换机的性能要远优于路由器，非常适用于数据交换频繁的局域网中；而路由器虽然路由功能非常强大，但数据包转发速率远低于三层交换机，更适合于数据交换不太频繁的不同类型网络的互联，如局域网与 Internet 互联。

课 后 习 题

1. 术语解释

网络互联　中继器　网桥　交换机　路由器　三层交换机

2. 网络互联有何实际意义？进行网络互联时，有哪些共同的问题需要解决？
3. 同样是物理层网络互联设备，中继器和集线器有什么异同？
4. 网卡的主要功能是什么？
5. 试对网桥和交换机的异同之处进行比较。
6. 试简述交换机地址学习的过程。
7. 试对路由器和三层交换机的异同之处进行比较。

第8章 传输层

学习目的

计算机网络本质活动是实现分布在不同地理位置主机之间的进程通信，以实现应用层的各种网络服务功能。本章讨论了传输层的基本功能，传输层向应用层提供的服务，以及实现这些服务的传输层 TCP 与 UDP 的基本内容。通过本章的学习，使读者掌握用户数据报协议(UDP)与传输控制协议(TCP)的基本内容，为读者进一步研究应用层与应用层协议打下基础。

学习要求

了解：网络环境中分布式进程通信的基本概念。

掌握：进程相互作用的 Client/Server 模型。

掌握：传输层的基本功能与服务质量。

掌握：用户数据报协议(UDP)的基本内容。

掌握：传输控制协议(TCP)的基本内容。

8.1 传输层功能概述

传输层是 OSI 参考模型的第 4 层，它为上一层提供了端到端(End to End)可靠的信息传递。物理层使得各链路上透明地传输比特流。数据链路层增强了物理层所提供的服务，使得相邻节点所构成的链路能够传输无差错的帧。网络层在数据链路层的基础上，提供路由选择、网络互联与协议转换等功能。而对于用户进程来说，希望得到的是端到端的服务(如主机 A 到主机 B 的 FTP)，传输层主要是建立应用间的端到端连接，并且为数据传输提供可靠或不可靠的连接服务。

8.1.1 传输层的基本功能

传输层是 OSI 参考模型中建立在网络层和会话层之间的一个层次，一般包括以下功能。

(1) 连接管理(Connection Management)：定义了允许两个用户像直接连接一样开始交谈的规则。通常把连接的定义和建立过程称为握手(Handshake)。传输层要建立、维持和终止一个会话，传输层与其对等系统建立面向连接服务的会话。在数据传输开始时，发送方和接收方的应用都要通知各自的操作系统初始化一个连接，一台主机发起的连接必须被另一台主机接收才行。当所有的同步操作完成之后，一个连接就建立了，数据传输也就开始了，在传输的过程中，两台主机还需要继续通过协议软件来通信以验证数据是否被正确接收。数据传输完成后，发送端主机发送一个标识数据传输结束的指示。接收端主机在数据

传输完成后确认数据传输结束，连接终止。

(2) 流量控制(Flow Control)：网络以一定的速度发送数据，从而防止网络拥塞造成数据报的丢失。传输层和数据链路层的流量控制区别在于：传输层定义了端到端用户之间的流量控制；数据链路层定义了两个中间的相邻节点的流量控制。

(3) 差错检测(Error Detection)：数据链路层的差错检测功能提供了可靠的链路传输，但无法保证源点和目的之间的传输完全无错，比如，网络中的路由器收到了完整无缺的 IP 分组，但是在将含有分组的帧重新格式化的过程中，出现了影响分组内容的错误。这种错误可能是由于软件或硬件问题导致路由器在进行分组期间引起的，并不是由于物理链路在进行数据传输过程中产生的差错，因此数据链路层的差错检测无法校验和识别出差错。传输层的差错检测机制会检测到这种类型的错误。

(4) 对用户请求的响应(Response to User's Request)：包括对发送数据和接收数据请求的响应，以及特定请求的响应，如用户可能要求高吞吐率、低延迟或可靠的服务。

(5) 建立无连接或面向连接的通信：TCP/IP 的 TCP 提供面向连接的传输层服务，UDP 则提供面向无连接的传输层服务。

传输层是 OSI 参考模型中最重要的一个层，其涉及在源主机与目的主机的进程之间提供端到端的可靠数据传输，并与当前使用的通信子网无关。为此，传输层引入了不少新概念与新机制。

首先，引入了网络进程标识的概念。在单机上，为了区别不同的进程，采用进程标识或进程号(Process ID)来唯一的标识进程。但是在网络环境中，这种由主机各自独立分配的进程号已经不能明确地标识进程了。在网络环境中，完整的进程标识需要这样的一种形式：源主机地址＋源进程标识，目的主机地址＋目的进程标识。

其次，传输层提供了一系列实现端到端进程之间可靠数据传输所必需的机制。包括面向连接服务的建立机制，即能够为高层数据的传输建立、维护与拆除传输连接，以实现透明、可靠的端到端的数据传输；端到端的错误恢复与流量控制，以能对网络层出现的丢包、乱序或重复等问题做出反应。

最后，当上层的协议数据包的长度超过网络层所能承载的最大数据传输单元时，传输层提供必要的分段功能，从而在接收方的对等层提供合并分段的功能。另外，有时候还要提供多路复用机制，即如果网络层提供的是面向连接服务，则传输层应该既能够将一个高层应用复用到多个网络层连接上，又能够将多个高层应用复用到一个网络层连接上。

总之，传输层通过扩展网络层服务功能，并通过传输层与高层之间的服务接口向高层提供了端到端节点之间的可靠数据传输，从而使系统之间实现高层资源的共享时不必再考虑数据通信方面的问题。传输层中完成相应功能的硬件与(或)软件被称为传输实体，其可能位于操作系统内核中、用户进程内、网络应用程序库中或网络接口卡上。在一个系统中，传输实体通过网络服务与其他对等的传输实体通信，进而向传输层用户(可以是应用进程，也可以是会话层协议)提供传输服务。

8.1.2 传输层的作用与地位

对于 TCP/IP 模型来说，传输层则是位于网络层和应用层之间的一个层次。那么为什

么还要在网络层之上再提供一个传输层而不直接面向会话层或应用层呢?

首先,传输层(又称运输层)在 TCP/IP 参考模型中位于互联网络层之上,它的功能是使源端和目的端主机上的对等实体可以进行会话(和 OSI 参考模型的传输层一样),该层定义了两个端到端的协议。

其次,从网络传输质量的角度,网络层虽然提供了从源网络到目的网络通信服务,但是其所提供的服务有可靠与不可靠之分,需要在网络层之上增加一个层次来弥补网络层所提供的服务质量的不足,以便为高层提供可靠的端到端通信。不可靠的 IP 协议提供“尽力而为”的服务,它不保证端到端数据传输的可靠性,IP 分组在传输过程中可能会出现丢包、乱序或重复。也可以这样理解,网络层及以下部分是由通信子网来完成的,由于历史及经济原因,通信子网往往是公用数据网,是资源子网中的端用户所不能直接控制的,用户不可能通过更换性能更好的路由器或增强数据链路层的纠错能力来提高网络层的服务质量,所以端用户只能依靠在自己主机上所增加的传输层来检测分组的丢失或数据的残缺并采取相应的补救措施。

从通信子网的角度,也可以这样理解,作为资源子网中的终端用户是不可能对通信子网内部加以直接控制的,即不可能通过更换性能更好的路由器或增强数据链路层的纠错能力来提高网络层的服务质量,只能依靠在自己主机上所增加的这个传输层来检测分组的丢失或数据的残缺并采取相应的补救措施。所以传输层不仅有存在的必要,它还是 OSI 参考模型中非常重要的一层,起到承上启下不可或缺的作用,从而被看成整个分层体系的核心。

但是,只有资源子网中的端设备才会具有传输层,通信子网中的设备一般至多只具备 OSI 下面 3 层的功能即通信功能。根据上述原因,通常又将 OSI 参考模型中的下面 3 层称为面向通信子网的层,而将传输层及以上的各层称为面向资源子网或主机的层。另一种划分则是将传输层及以下的各层统称为面向数据通信的层,而将传输层之上的会话层、表示层及应用层这些不包含任何数据传输功能的层统称为面向应用的层,如图 8.1 所示。

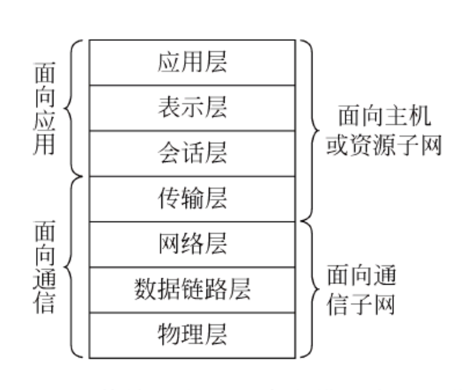


图 8.1 传输层在 OSI 参考模型中的地位

8.1.3 网络服务与服务质量

服务是计算机网络中一个非常重要的概念,它描述了计算机网络体系结构中相邻层之间的关系。在计算机网络层次结构中,N 层总是利用 $N-1$ 层所提供的服务,向 $N+1$ 层提供更加完善和更高质量的服务。N 层是服务的提供者,$N+1$ 层是服务的用户。例如,数据链路层利用了物理层所提供的原始比特流服务,向网络层提供了相邻节点间的可靠数据传输服务。作为服务,必然涉及服务质量的问题。在网络层次结构的每一层上都有服务质量的问题。例如,在物理层,提供的原始比特流传输有速率高低、误码率高低之分;在数据链路层,所提供的相邻节点间的帧传输服务有面向连接与面向无连接之分;在网络层,源到目的的分组传输根据通信子网的不同被分成了虚电路与数据报服务;在 IP 数据报中,还通过

报头中的 ToS 字段提供了不同类型与优先级的数据报传输服务。在计算机网络中，通常将这种服务质量简称为 QoS。

那为什么到了传输层才将 QoS 明确地提出来呢？首先，传输层是 OSI 参考模型中面向通信的最高层，在它上面的各层都是面向应用的层，因此在传输层衡量网络通信提供给高层应用的服务质量是最为恰当的。其次，尽管传输层所提供的很多 QoS 指标是由通信子网的服务和服务质量所决定的，但传输层还是可以通过许多机制来改善网络服务的可靠性。

衡量传输层服务质量的主要参数如下。

(1) 连接建立延迟：从传输服务用户要求建立连接到收到连接确认之间所经历的时间；它包括了远端传输实体的处理延迟；连接建立延迟越短，服务质量越好。

(2) 连接建立失败的概率：在最大连接建立延迟时间内，连接未能建立的可能性；由于网络拥塞，缺少缓冲区或其他原因造成的失败。

(3) 吞吐率：吞吐率是在某个时间间隔内测得的每秒钟传输的用户数据的字节数；每个传输方向分别用各自的吞吐率来衡量。

(4) 传输延迟：传输延迟是指从源主机传输用户发送报文开始到目的主机传输用户接收到报文为止的时间；每个方向的传输延迟是不同的。

(5) 残余误码率：残余误码率用于测量丢失或乱序的报文数占整个发送的报文数的百分比；理论上残余误码率应为零，实际上它可能是一较小的值。

(6) 安全保护：安全保护为传输用户提供了传输层的保护，以防止未经授权的第三方读取或修改数据。

(7) 优先级：为传输用户提供用以表明哪些连接更为重要的方法；当发生拥塞事件时，确保高优先级的连接先获得服务。

(8) 恢复功能：当出现内部问题或拥塞情况下，传输层本身自发终止连接的可能性。

传输层提供的服务也有面向连接与面向无连接之分。对于面向连接服务，可以在连接建立阶段，就各种服务参数指定希望得到的级别和可以接受的最低限度，并就此与目的主机进行必要的协商。

请注意并不是所有的传输层连接都要求提供上述所有的服务参数，决定于传输层用户的具体需求。当用户指定相应的服务参数，而传输层通过检查服务质量参数发现自己不能达到用户所希望的质量参数时，它可以采用两种方式：一是决定不与目的主机建立传输连接，并直接通知传输用户连接请求失败；二是降低要求，然后再请求建立连接。

8.2 TCP

由于 TCP/IP 的网络层提供的是面向无连接的数据报服务，也就是说，IP 数据报传输会出现丢失、重复或乱序的情况，因此在 TCP/IP 网络中传输层就变得极为重要。TCP/IP 的传输层提供了两个主要的协议即传输控制协议(Transport Control Protocol，TCP)和用户数据报协议(User Datagram Protocol，UDP)。

尽管 TCP/IP 的网络层提供的是一种面向无连接的 IP 数据报服务，但传输层的 TCP 旨在向 TCP/IP 的应用层提供一种端到端的面向连接的可靠的数据流传输服务。TCP 常

用于一次传输要交换大量报文的情形,如文件传输、远程登录等。

为了实现这种端到端的可靠传输,TCP 必须规定传输层的连接建立与拆除的方式、数据传输格式、确认的方式、目标应用进程的识别以及差错控制和流量控制机制等。与所有网络协议类似,TCP 将自己所要实现的功能集中体现在了 TCP 的协议数据单元中。

8.2.1 TCP 分段的格式

TCP 的协议数据单元被称为分段(Segment),TCP 通过分段的交互来建立连接、传输数据、发出确认、进行差错控制、流量控制及关闭连接。分段分为两部分,即分段头和数据,所谓分段头就是 TCP 为了实现端到端可靠传输所加上的控制信息,而数据则是指由高层即应用层来的数据。图 8.2 给出了 TCP 分段头的格式,其中有关字段的说明如下。

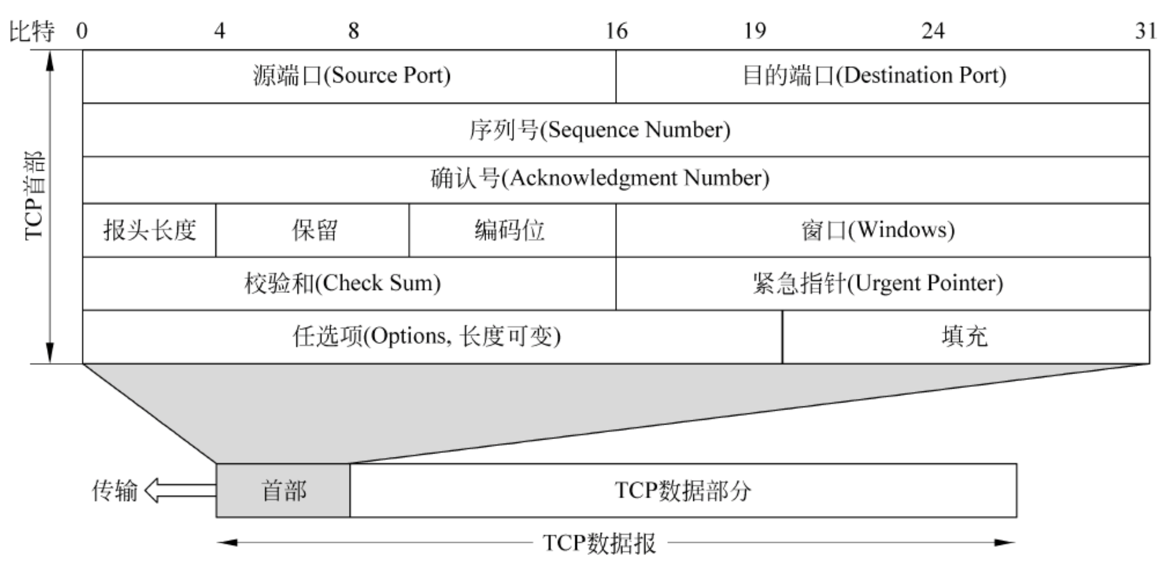


图 8.2 TCP 报文段的首部

(1) 源端口:占 16b,分段的源端口号。

(2) 目的端口:占 16b,分段的目的端口号。

(3) 序列号:占 32b,分段的序列号,表示该分段在发送方的数据流中的位置,用来保证到达数据顺序的编号。

(4) 确认号:占 32b,下一个期望接收的 TCP 分段号,相当于是对对方所发送的并已被本方所正确接收的分段的确认。序列号和确认号共同用于 TCP 服务中的确认、差错控制。

(5) 报头长度:TCP 头长,以 32b 字长为单位。实际上相当于给出数据在数据段中的开始位置。

(6) 保留:占 6b,为将来的应用而保留,目前置为 0。

(7) 编码位:占 6b,TCP 分段有多种应用,如建立或关闭连接、传输数据、携带确认等,这些编码位用于给出与分段的作用及处理有关的控制信息,详细参见表 8.1。

(8) 窗口:占 16b,窗口的大小表示发送方可以接收的数据量,单位为字节。使用可变大小的滑动窗口协议来进行流量控制。

(9) 校验和:占 16b,用于对分段首部和数据进行校验。通过将所有 16b 以补码形式相

加，然后再对相加和取补，正常情况下应为0。

(10) 紧急指针：占16b，给出从当前序列号到紧急数据位置的偏移量。

(11) 任选项：长度可变。TCP只规定了一种选项，即最大报文段长度(MSS)。

(12) 填充：当任选项字段长度不足32b字长时，需要加以填充。

(13) 数据：来自高层即应用层的协议数据。

表8.1 TCP分段头中的编码位字段的含义

编码位(从左到右)的标识	该位置1的含义
紧急比特(URG)	表示启用了紧急指针字段
确认比特(ACK)	表示确认字段是有效的
推送比特(PSH)	请求急迫操作，即分段一到马上发送应用程序而不等到接收缓冲区满时才发送应用程序
复位比特(RST)	连接复位。复位因主机崩溃或其他原因而出现错误的连接，也可用于拒绝非法的分段或拒绝连接请求
同步比特(SYN)	与ACK合用以建立TCP连接。如SYN=1，ACK=0表示连接请求；而SYN=1，ACK=1表示同意建立连接
终止比特(FIN)	表示发送方已无数据要发送从而要释放连接，但接收方仍可继续接收发送方此前发送的数据

8.2.2 端口和套接字

上面TCP分段的格式中出现了“源端口”和“目的端口”字段，“端口”是英文Port的意译，作为计算机术语，“端口”被认为是计算机与外界通信交流的出入口。

由网络OSI参考模型可知，传输层与网络层最大的区别是传输层提供进程通信能力，网络通信的最终地址不仅包括主机地址，还包括可描述网络进程的某种标识。所以TCP/IP所涉及的端口是指用于实现面向连接或无连接服务的通信协议端口，是对网络通信进程的一种标识，其属于一种抽象的软件结构，包括一些数据结构和I/O(输入/输出)缓冲区，故属于软件端口范畴。应用程序(调入内存运行后一般称为进程)通过系统调用与某传输层端口建立绑定(Binding)后，传输层传给该端口的所有数据都被建立这种绑定的相应进程所接收，相应进程发给传输层的数据也都从该端口输出。在TCP/IP的实现中，端口操作类似于一般的I/O操作，进程获取一个端口，相当于获取本地唯一的I/O文件，可以用一般的读/写方式访问。

每个端口都拥有一个叫作端口号的整数描述符，用来标识不同的端口或进程。在TCP/IP传输层，定义一个16b长度的整数作为端口标识，也就是说可定义2^{16}个端口，其端口号从0到$2^{16}-1$。由于TCP/IP传输层的TCP和UDP两个协议是两个完全独立的软件模块，因此各自的端口号也相互独立，即各自可独立拥有2^{16}个端口。

正如图8.3所示，每种应用层协议或应用程序都具有与传输层唯一连接的端口，并且使用唯一的端口号将这些端口区分开来。当数据流从某一个应用发送到远程网络设备的某一个应用时，传输层根据这些端口号，就能够判断出数据是来自哪一个应用，想要访问另一台网络设备的哪一个应用，从而将数据传输到相应的应用层协议或应用程序。

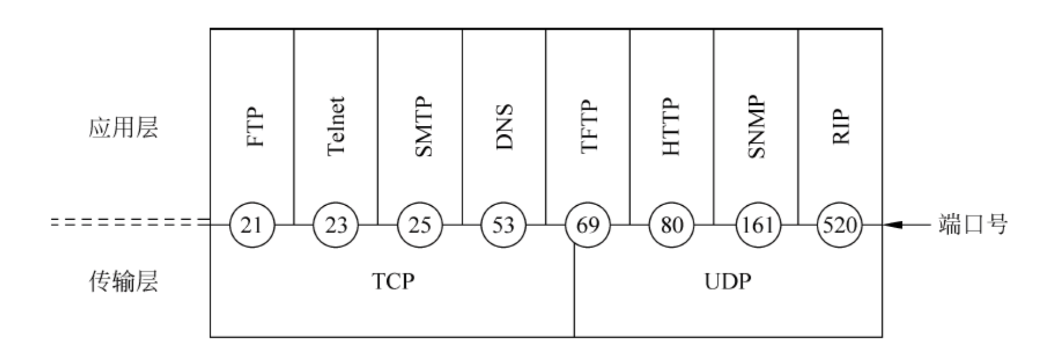


图 8.3　应用层与传输层之间的接口

端口根据其对应的协议或应用不同,被分配了不同的端口号。负责分配端口号的机构是 Internet 编号管理局(IANA)。目前,端口的分配有 3 种情况,这 3 种不同的端口可以根据端口号加以区别。

(1) 保留端口。这种端口号一般都小于 1024。它们基本上都被分配给了已知的应用协议(如图 8.3 中的部分端口)。目前,这一类端口的端口号分配已经被广大网络应用者接受,形成了标准,在各种网络的应用中调用这些端口号就意味着使用它们所代表的应用协议。这些端口由于已经有了固定的使用者,所以不能被动态地分配给其他应用程序。表 8.2 给出了一些常用的保留端口。

(2) 动态分配的端口。这种端口的端口号一般都大于 1024。这一类的端口没有固定的使用者,它们可以被动态地分配给应用程序使用。也就是说,在使用应用软件访问网络时,应用软件可以向系统申请一个大于 1024 的端口号临时代表这个软件与传输层交换数据,并且使用这个临时的端口与网络上的其他主机通信。

(3) 注册端口。注册端口比较特殊,它也是固定为某个应用服务的端口,但是它所代表的不是已经形成标准的应用层协议,而是某个软件厂商开发的应用程序。

表 8.2　TCP 和 UDP 的一些常用保留端口

	端口号	关键字	应用协议
UDP 保留端口举例	53	DNS	域名服务系统
	69	TFTP	简单文件传输协议
	161	SNMP	简单网络管理协议
	520	RIP	RIP 路由选择协议
TCP 保留端口举例	21	FTP	文件传输协议
	23	Telnet	虚拟终端协议
	28	SMTP	简单邮件传输协议
	53	DNS	域名服务
	80	HTTP	超文本传输协议

某些软件厂商通过使用注册端口,使它的特定软件享有固定的端口号,而不用向系统申请动态分配的端口号。通常,这些特定的软件要使用注册端口,其厂商必须向端口的管理机构注册。

大多数注册端口的端口号大于1024。

TCP和UDP都允许16b的端口值，分别能够提供65 536个端口。不论端口号大于还是小于1024，以上3种端口都分别属于TCP和UDP。当然，也有些协议的端口既属于TCP也属于UDP。

当网络中的两台主机进行通信时，为了表明数据是由源端的哪一种应用发出的，以及数据所要访问的是目的端的哪一种服务，TCP/IP会在传输层封装数据段时，把发出数据的应用程序的端口作为源端口，把接收数据的应用程序的端口作为目的端口，添加到数据段的头中，从而使主机能够同时维持多个会话的连接，使不同应用程序的数据不发生混淆。一台主机上的多个应用程序可同时与其他多台主机上的多个对等进程进行通信，所以需要对不同的虚电路进行标识。对TCP虚电路连接采用发送端和接收端的套接字(Socket)组合来识别，如Socket1、Socket2。所谓套接字，实际上是一个通信端点，每个套接字都有一个套接字序号，包括主机的IP地址与一个16b的主机端口号，如主机IP地址、端口号。图8.4表现了源端口与目的端口的作用。

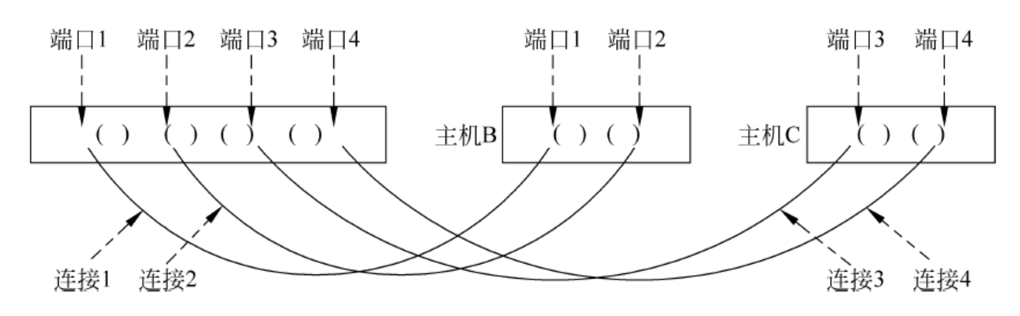


图8.4 端口的概念示意图

应该指出，尽管采用了上述的端口分配模式，但在实际使用中，经常会采用端口重定向技术。所谓端口重定向，是指将一个著名端口重定向到另一个端口，例如默认的HTTP端口是80，不少人将它重定向到另一个端口，如8080。

端口在传输层的作用有点类似IP地址在网络层的作用或MAC地址在数据链路层的作用，只不过IP地址和MAC地址标识的是主机，而端口标识的是网络应用进程。由于同一时刻一台主机上会有大量的网络应用进程在运行，所以需要有大量的端口号来标识不同的进程。

正是由于TCP使用通信端点来识别连接，才使得一台计算机上的某个TCP端口号可以被多个连接所共享，从而程序员可以设计出能同时为多个连接提供服务的程序，而不需要为每个连接设置各自的本地端口号。

8.2.3 TCP的连接建立和拆除

TCP连接包括建立与拆除两个过程。TCP使用三次握手协议来建立连接。连接可以由任何一方发起，也可以由双方同时发起。一旦一台主机上的TCP软件已经主动发起连接请求，运行在另一台主机上的TCP软件就被动地等待握手。图8.5给出了三次握手建立TCP连接的简单示意。主机1首先发起TCP连接请求，并在所发送的分段中将编码位字段中的SYN位被置为1、ACK位被置为0。主机2收到该分段，若同意建立连接，则发送一

个连接接收的应答分段,其中编码位字段的 SYN 和 ACK 位均被置为 1,指示对第一个 SYN 报文段的确认,以继续握手操作;否则,主机 2 要发送一个将 RST 位被置为 1 的应答分段,表示拒绝建立连接。主机 1 收到主机 2 发来的同意建立连接分段后,还有再次进行选择的机会,若其确认要建立这个连接,则向主机 2 发送确认分段,用来通知主机 2 双方已完成建立连接;若其不想建立这个连接,则可以发送一个将 RST 位被置为 1 的应答分段来告知主机 2 拒绝建立连接。

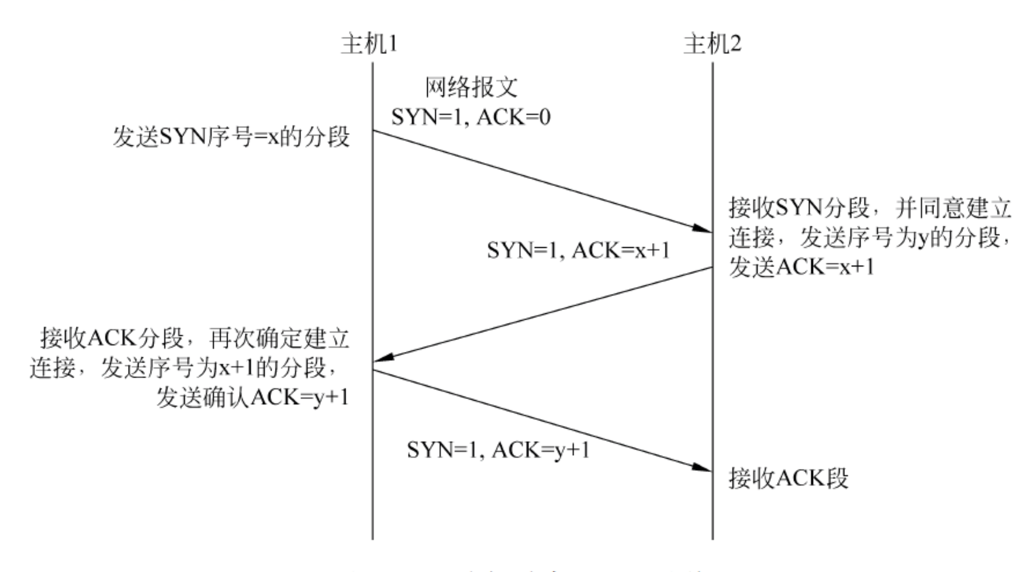


图 8.5　三次握手建立 TCP 连接

不管是哪一方先发起连接请求,一旦连接建立,就可以实现全双工的数据传输,而不存在主从关系。TCP 将数据流看作字节的序列,将从用户进程接收的任意长的数据,分成不超过 64 KB(包括 TCP 头在内)的分段,以适合 IP 数据报的载荷能力。所以对于一次传输要交换大量报文的应用(如文件传输、远程登录等),往往需要以多个分段进行传输。

数据传输完成后,还要进行 TCP 连接的拆除或关闭。TCP 协议使用修改的三次握手协议来关闭连接,以结束会话。TCP 连接是全双工的,可以看作两个不同方向的单工数据流传输。所以一个完整连接的拆除涉及两个单向连接的拆除。如图 8.6 所示,当主机 1 的 TCP 数据已发送完毕时,在等待确认的同时可发送一个将编码位字段的 FIN 位被置 1 的分段给主机 2,若主机 2 已正确接收主机 1 的所有分段,则会发送一个数据分段正确接收的确认分段,同时通知本地相应的应用程序,对方要求关闭连接,接着再发送一个对主机 1 所发送的 FIN 分段进行确认的分段。否则,主机 1 就要重传那些主机 2 未能正确接收的分段。收到主机 2 关于 FIN 确认后的主机 1 需要再次发送一个确认拆除连接的分段,主机 2 收到该确认分段意味着从主机 1 到主机 2 的单向连接已经结束。但是,此时在相反方向上,主机 2 仍然可以向主机 1 发送数据,直到主机 2 数据发送完毕并要求关闭连接。一旦当两个单向连接都被关闭,则两个端节点上的 TCP 软件就要删除与这个连接的有关记录,于是原来所建立的 TCP 连接被完全释放。

8.2.4　TCP 可靠数据传输技术

TCP 采用了许多与数据链路层类似的机制来保证可靠的数据传输,如采用序列号、确

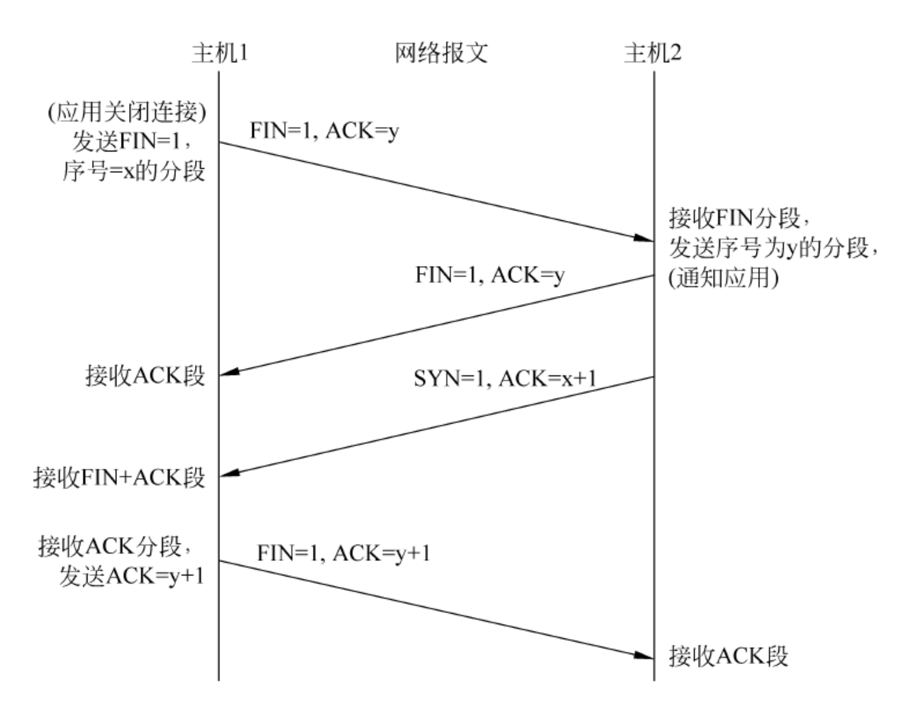


图 8.6 用于关闭连接的四次握手操作

认、滑动窗口协议等。只不过 TCP 的目的是为了实现端到端节点之间的可靠数据传输,而数据链路层协议则为了实现相邻节点之间的可靠数据传输。

首先,TCP 要为所发送的每一个分段加上序列号,保证每一个分段能被接收方接收,并只被正确地接收一次。

其次,TCP 采用具有重传功能的积极确认技术作为可靠数据流传输服务的基础。这里,"确认"是指接收端在正确收到分段之后向发送端回送一个确认(ACK)信息。发送方将每个已发送的分段备份在自己的发送缓冲区里,而且在收到相应的确认之前是不会丢弃所保存的分段的。"积极"是指发送方在每一个分段发送完毕的同时启动一个定时器,假如定时器的定时期满而关于分段的确认信息尚未到达,则发送方认为该分段已丢失并主动重发。为了避免由于网络延迟引起迟到的确认和重复的确认,TCP 规定在确认信息中捎带一个分段的序号,使接收方能正确地将分段与确认联系起来。

最后,采用可变长的滑动窗口协议进行流量控制,以防止由于发送端与接收端之间的不匹配而引起数据丢失。这里所采用的滑动窗口协议与数据链路层的滑动窗口协议在工作原理上是完全相同的,唯一的区别在于滑动窗口协议用于传输层是为了在端到端节点之间实现流量控制,而用于数据链路层是为了在相邻节点之间实现流量控制。TCP 采用可变长的滑动窗口,使得发送端与接收端可根据自己的 CPU 和数据缓存资源对数据发送和接收能力做出动态调整,从而灵活性更强,也更合理。例如,假设主机 1 有一个大小为 4096B 长的缓冲区,向主机 2 发送 2048B 长度的数据分段,则在未收到主机 2 的关于该 2048B 长度分段的确认之前,主机 1 向其他主机只能声明自己有一个 2048B 长度的发送缓冲区。过了一段时间后,假定主机 1 收到了来自主机 2 的确认,但其中声明的窗口大小为 0,这表明主机 2 虽然已经正确收到主机 1 前面所发送的分段,但目前主机 2 已不能接收任何来自主机 1 的新的分段了,除非以后主机 2 给出窗口大于 0 的新信息。

8.2.5　TCP 流量控制

TCP 采用大小可变的滑动窗口机制实现流量控制功能。窗口的大小是字节。在 TCP 报文段首部的窗口字段写入的数值就是当前给对方设置发送窗口的数据的上限。

在数据传输过程中，TCP 提供了一种基于滑动窗口协议的流量控制机制，用接收端接收能力(缓冲区的容量)的大小来控制发送端发送的数据量。

在建立连接时，通信双方使用 SYN 报文段或 ACK 报文段中的窗口字段捎带着各自的接收窗口尺寸，即通知对方从而确定对方发送窗口的上限。在数据传输过程中，发送方按接收方通知的窗口尺寸和序号发送一定量的数据，接收方根据接收缓冲区的使用情况动态调整接收窗口尺寸，并在发送 TCP 报文段或确认段时捎带新的窗口尺寸和确认号通知发送方。

如图 8.7 所示，设主机 A 向主机 B 发送数据。双方确定的窗口值是 400。设一个报文段为 100B 长，序号的初始值为 1(即 SEQ=1)。在图 8.7 中，主机 B 进行了 3 次流量控制。第一次将窗口减小为 300B，第二次将窗口又减为 200B，最后一次减至零，即不允许对方再发送数据了。这种暂停状态将持续到主机 B 重新发出一个新的窗口值为止。

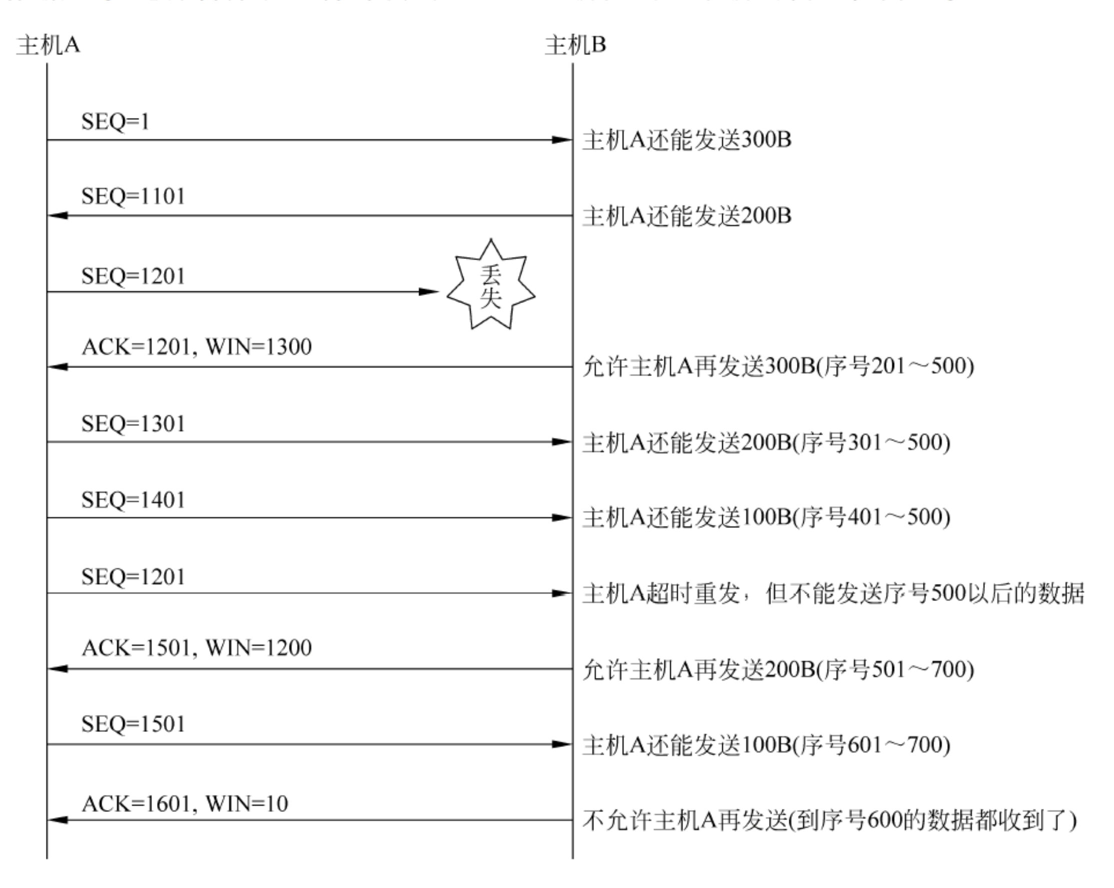


图 8.7　利用可变滑动窗口进行流量控制

在以太网的环境下，当发送端不知道对方窗口大小时，便直接向网络发送多个报文段，直至收到对方通告的窗口大小为止。但如果在发送方和接收方有多台路由器和较慢的链路时，就可能出现一些问题，一些中间路由器必须缓存分组，并有可能耗尽存储空间，这样就会

严重降低 TCP 连接的吞吐量。这时采用了一种称为慢启动的算法，慢启动为发送方的 TCP 增加一个拥塞窗口，当与另一个网络的主机建立 TCP 连接时，拥塞窗口被初始化为 1 个报文段(即另一端通告的报文段大小)，每收到一个 ACK，拥塞窗口就增加一个报文段(以字节为单位)。发送端取拥塞窗口与通告窗口中的最小值作为发送上限。拥塞窗口是发送方使用的流量控制，而通告窗口则是接收方使用的流量控制。开始时发送一个报文段，然后等待 ACK。当收到该 ACK 时，拥塞窗口从 1 增加为 2，即可发送两个报文段。当收到这两个报文段的 ACK 时，拥塞窗口就增加为 4。这是一种指数增加的关系。在某些互联网的中间某些点上可能达到了互联网的容量，于是中间路由器开始丢弃分组，这时通知发送方它的拥塞窗口开得过大。

8.3 UDP

8.3.1 UDP 概述

与 TCP 相反，UDP 提供的是不可靠的、无连接的数据传输服务。UDP 在完成进程之间的通信过程中，只提供了非常有限的差错检验功能，不提供数据接收的确认、排序和流量控制等功能，因此数据传输可能会出现丢失、乱序和重复现象。从这一点看，UDP 与网络层的 IP 类似，所以被称为用户数据报协议。

由于 UDP 的功能简单，所以协议的设计也相对简单。提供 UDP 这么一个较简单的传输层协议是希望以较小的开销(Overhead)来实现网络进程间的通信。UDP 非常适用于通信子网的传输质量或可靠性较高的网络环境，如局域网。另外，对于那些一次性传输数据量较小同时对数据传输可靠性要求又不高的网络应用，例如，SNMP、DNS、TFTP 数据的传输，也可以采用 UDP。因为对于这些一次性传输数据量较小的网络应用，若采用 TCP 服务，则所付出的关于连接建立、维护和拆除的开销是非常不合算的。近年来，随着 IP 电话、视频会议、流媒体通信、网络多播等实时应用的流行，UDP 也被用来作为这些应用的传输层协议。这类应用有一些共同的特点：要求源主机提供恒定的数据发送速率；在网络出现拥塞时，允许丢失部分数据；网络延迟要尽可能地小。因此，UDP 能够很好地适应它们的应用需求。使用 UDP 为传输层协议的网络应用其可靠性的问题需要由使用 UDP 的应用程序来解决。

UDP 只在 IP 的数据报服务之上增加了很少的功能，这就是端口的功能(有了端口，运输层就能进行复用和分用)和差错检测的功能。UDP 在某些方面有其如下特殊的优点。

(1) 发送数据之前不需要建立连接，减少了开销和发送数据之前的时延。

(2) UDP 不使用拥塞控制，也不保证可靠交付，因此主机不需要维持具有许多参数的、复杂的连接状态表。

(3) UDP 用户数据报只有 8B 的首部开销。

(4) 由于 UDP 没有拥塞控制，因此网络出现的拥塞不会使源主机的发送速率降低，这对某些实时应用是很重要的。很多的实时应用(如 IP 电话、实时视频会议等)要求源主机以恒定的速率发送数据，并且允许在网络发生拥塞时丢失一些数据，但却不允许数据有太大的

时延,UDP 正好适合这种要求。

UDP 常用于一次性传输数据量较小的网络应用,如 SNMP、DNS 应用数据的传输。因为对于这些一次性传输数据量较小的网络应用,若采用 TCP 服务,则所付出的关于连接建立、维护和拆除的开销是非常不合算的。

8.3.2 UDP 数据报的首部格式

UDP 有两个字段:数据字段和首部字段。首部字段只有 8B,由 4 个字段组成,每个字段都是 2B,如图 8.8 所示,各字段意义如下。

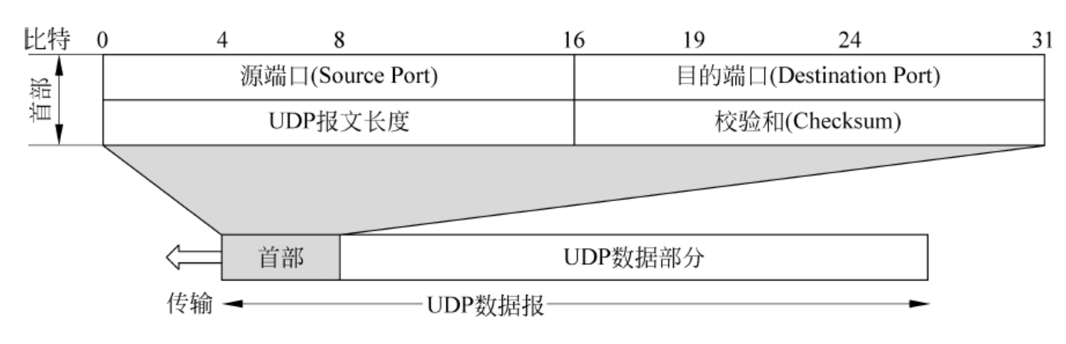


图 8.8 UDP 数据报的首部

一个 UDP 数据报由 UDP 报头和 UDP 数据两部分组成。其中,UDP 报头的固定长度为 8B,由 4 个各 16b 长度的字段组成。它们分别为 UDP 源端口、UDP 目的端口、UDP 数据报长度以及校验和。

(1) 源端口:占 16b,源端口号。

(2) 目的端口:占 16b,目的端口号。

(3) UDP 报文长度:占 16b,UDP 用户数据报的长度。

(4) 校验和:占 16b,防止 UDP 用户数据报在传输中出错。

在 UDP 中也采用与 TCP 中类似的端口概念来标识同一主机上的不同网络进程,并且两者在分配方式上也是类似的。UDP 与应用层之间的端口都是用报文队列来实现的。

8.3.3 UDP 的工作过程

UDP 提供无连接服务,用户数据报在发送之前不需要建立连接。当应用进程有报文需要通过 UDP 发送时,它将此报文直接交给执行 UDP 的传输层实体。报文的长度要足够短,以便能装入一个 UDP 数据报中,所以只有发送短报文的进程才选用 UDP。UDP 传输层实体在得到应用进程的报文后,为它加上 UDP 报头,变成 UDP 数据报后交给网络层。网络层在 UDP 用户数据报前面加上 IP 报头,形成 IP 分组,再交给数据链路层。数据链路层在 IP 分组上加上帧头和帧尾,变成一个帧,然后通过物理层发送出去。对于目的端,则是一个相反的拆封过程。

由于 UDP 提供无连接服务,所以每个 UDP 用户数据报的传输路径都是独立的。即使那些 UDP 用户数据报的源端口号和目的端口号相同,它们在网络上的传输路径也可能是

不同的，取决于网络层为每个数据报所进行的路径选择。一个先发送的 UDP 用户数据报因为网络路径的不同，可能会较一个晚发送的 UDP 用户数据报后到。

UDP 是一个不可靠的协议，不提供确认、流量控制等可靠传输机制，所以对于 UDP 的接收端来说，一旦当到来的报文过多时，就会因为溢出而使报文丢失。另外，由于 UDP 只提供简单的校验和，没有确认、重传等差错控制机制，因此当接收进程通过校验和发现传输出错时，只是简单地将该出错的用户数据报丢弃，并不向发送进程提供错误通知。相应地，采用 UDP 的应用进程需要在应用层提供必要的差错控制机制。

为了区分同一台主机并发运行的多个 UDP 进程，传输层实体采用了一种与 UDP 端口相关联的用户数据报传输队列机制。图 8.9 给出了一对用户进程通过 UDP 进行数据交换时，用户数据报传输队列工作原理的简单示意。

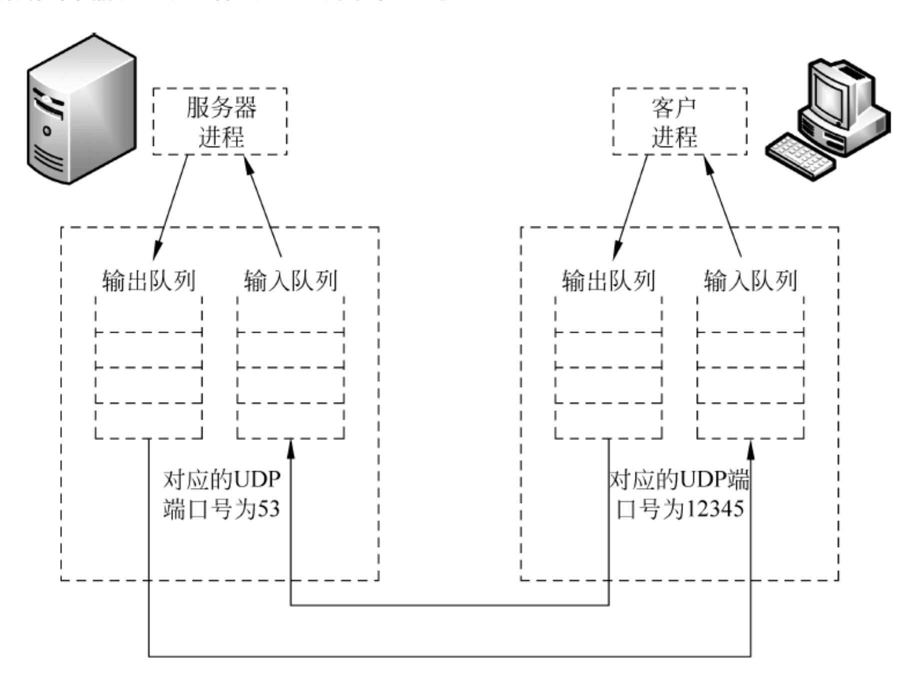


图 8.9 UDP 用户数据报传输队列的工作原理

当客户进程启动时，UDP 为该进程分配一个临时端口号（假定为 12345），并同时创建与该端口号对应的一个输出队列和一个输入队列。所有该客户进程要发送的用户数据报，被写入输出队列；而从服务器端对等进程返回的用户数据报，则放在该客户进程端口号所对应的输入队列中。如果输入队列产生溢出或创建问题时，客户端将无法接收从服务器端对等进程所返回的数据，此时，客户端会丢弃这些用户数据报，并请求客户机通过 ICMP 向服务器端发送“端口不可到达”的出错报文。如果输出队列发生溢出时，操作系统就会要求客户进程降低用户数据报的发送速率。

在服务器端，用户数据报传输队列的创建机制是不同的。只要服务器进程开始运行，UDP 进程就会用相应的著名端口号去创建一个输入队列和一个输出队列。只要服务器进程在运行，这些队伍就一直存在，不管是否有客户进程在请求。当客户的 UDP 请求到达时，服务器端的 UDP 要检查对应于该用户数据报目的端口的输入队列是否已经存在，若已经存在，则将收到的客户 UDP 请求放在该输入队列的末尾。否则，就丢弃该用户数据报，并通过 ICMP 向客户端发送“端口不可到达”的报文。对于服务器进程而言，不管 UDP 请

求是否来自不同的客户端，都要被放入同一个输入队列。当输入队列发生溢出时，UDP服务器进程就丢弃该用户数据报，并请求通过ICMP向客户端发送“端口不可到达”的报文。当服务器进程需要向客户发送用户数据报时，它就将发送报文放到该服务器进程端口号所对应的输出队列。若输出队列发生溢出，操作系统会要求该服务器进程在继续发送报文之前先等待一段时间。

课后习题

1. 术语解释

服务质量(QoS)　TCP　UDP　端口号套接字　分段　用户数据报　自由端口

2. 请说明对OSI传输层重要性的理解。

3. TCP/IP的传输层为什么要设计两个提供不同服务的协议？

4. 什么是端口？它在传输层的作用是什么？

5. 请说明传输层TCP采用了哪些机制来保证端到端节点之间的可靠传输？

6. 请列举5～10个你所知道的著名的TCP端口或UDP端口，并说明它们是提供什么网络应用的。

7. 试说明传输层的作用。网络层提供数据报或虚电路服务对上面的传输层有何影响？

8. 试用具体例子说明为什么在传输连接建立时要使用三次握手。说明如不这样做可能会出现什么情况。

9. 为什么在TCP首部中有一个首部长度字段，而UDP的首部中就没有这个字段？

第9章 Internet技术与应用层

学习目的

通过本章的学习，帮助读者掌握域名系统（DNS）、文件传输协议（FTP）、电子邮件（E-mail）、WWW 服务等应用层服务的工作原理与应用层协议的基本概念，深入理解网络协议的层次结构、客户/服务器的交互过程等。为进一步学习网络应用技术与网络编程技术打下坚实的基础。

学习要求

了解：TCP/IP 协议栈与应用层协议之间的关系。

掌握：域名系统的基本工作原理。

掌握：电子邮件的基本工作原理。

掌握：文件传输协议的基本工作原理。

掌握：WWW 服务的基本工作原理。

掌握：应用层协议的分析方法。

Internet 是一个全球性的巨大的计算机网络体系，它把全球数百万个计算机网络，数亿台计算机主机连接起来，包含了难以计数的信息资源，向全世界提供信息服务。从网络通信技术的观点来看，Internet 是一个以 TCP/IP 通信协议为基础，连接各个国家、各个部门、各个机构计算机网络的数据通信网。从信息资源的观点来看，Internet 是一个将各个领域、各个学科的各种信息资源融为一体、供网上用户共享的数据资源网。

一般认为，Internet 的定义至少包含以下 3 个方面的内容。

（1）Internet 是一个基于 TCP/IP 协议簇的国际互联网络。

（2）Internet 是一个网络用户的团体，用户使用网络资源，同时也为该网络的发展壮大贡献力量。

（3）Internet 是所有可被访问和利用的信息资源的集合。

9.1 Internet 概 述

Internet 起源于美国，1969 年实现的 ARPAnet（Advanced Research Project Agency network）是 Internet 的前身。1986 年在美国国家科学基金会 NSF 的资助下，使用 TCP/IP 的 NSFNET 开始建设，它鼓励各地区网络吸收非学术的商业用户，并最终取代了 ARPAnet 成为 Internet 的骨干网。

随着万维网（World Wide Web，WWW）在 Internet 上被广泛使用，使广大非网络专业人员也能方便地使用网络，这成为 Internet 指数级增长的主要驱动力。

1987 年 9 月 14 日，北京计算机应用技术研究所发出了中国第一封电子邮件：“Across the Great Wall We Can Reach Every Corner in the World.”(越过长城，走向世界)，揭开了中国人使用互联网的序幕。

9.1.1 Internet 在我国的发展

1988 年 7 月，中国科学院高能物理研究所采用 X.25 协议使该单位的 DECnet 成为西欧中心 DECnet 的延伸，实现了计算机国际远程联网以及与欧洲和北美地区的电子邮件通信。

1990 年 11 月 28 日，钱天白教授代表中国正式在 SRI-NIC(Stanford Research Institute's-Network Information Center)注册登记了中国的顶级域名 CN，从此中国的网络有了自己的身份标识。

1993 年 3 月 2 日，中国科学院高能物理研究所连入美国斯坦福大学线性加速器中心(SLAC)的 64Kbps 专线正式开通。这条专线仍是中国部分连入 Internet 的第一根专线。

1994 年 4 月初，美国国家科学基金会 NSF 同意了 NCFC(The National Computing and Networking Facility of China，中国国家计算与网络设施)正式连入 Internet 的要求。

1994 年 4 月 20 日，NCFC 工程连入 Internet 的 64K 国际专线开通，实现了与 Internet 的全功能连接。从此中国被国际上正式承认为真正拥有全功能 Internet 的第 77 个国家。

1994 年 5 月 21 日，中国科学院计算机网络信息中心完成了中国国家顶级域名 CN 服务器的设置，改变了中国的 CN 顶级域名服务器一直放在国外的历史。从此，中国的 Internet 开始了迅速发展的时期，在此期间相继建设了四大 Internet 骨干网，开启了铺设中国信息高速公路的历程。这四大骨干网分别是中国科技网、中国金桥信息网、中国公用计算机互联网、中国教育和科研计算机网。

根据 CNNIC 的《第 22 次中国互联网络发展状况统计报告》，截至 2008 年 6 月底，我国互联网用户达到 2.53 亿人，跃居世界第一位。

9.1.2 Internet 的相关机构

Internet 最大的特点是管理上的开放性，它没有集中的管理机构，但为了促进 Internet 运行所需的标准兼容性并确保 Internet 的持续发展，先后成立了一些机构自愿承担必需的管理职责，并且遵循自下而上的结构原则。

1. Internet 协会及其组织机构

Internet 协会(Internet Society，ISOC)成立于 1992 年，总部在美国，该协会是一个推动、支持和促进 Internet 不断增长和发展的专业组织，它把 Internet 作为全球研究通信的基础设施。ISOC 的网址是 www.isoc.org，ISOC 的组织机构如图 9.1 所示。

1) Internet 体系结构委员会

Internet 体系结构委员会(Internet Architecture Board，IAB)创建于 1992 年 6 月，是 ISOC 的技术咨询机构，其职能是负责 Internet 标准的最后编辑和技术审核。

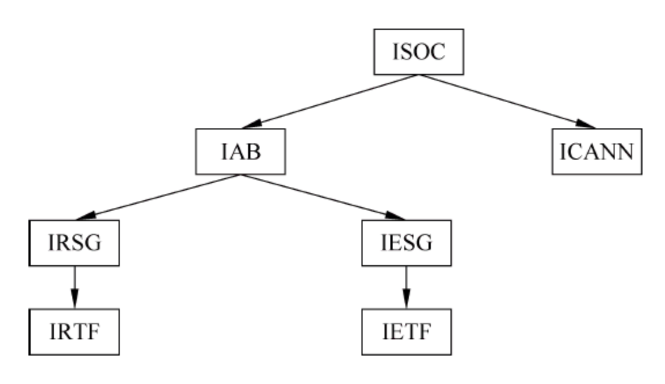


图 9.1 ISOC 的组织机构

2）互联网名称与数字地址分配机构(ICANN)

ICANN 成立于 1998 年 10 月，是一个集合了全球网络界商业、技术及学术各领域专家的非营利性国际组织，负责互联网协议(IP)地址的空间分配、协议标识符的指派、通用顶级域名以及国家和地区顶级域名系统的管理，以及根服务器系统的管理。

3）Internet 网络信息中心(InterNIC)

InterNIC 负责 Internet 域名注册和域名数据库的管理。InterNIC 网站目前由 ICANN 负责维护，提供互联网域名登记服务的公开信息。

4）Internet 工程任务组(IETF)

Internet 工程任务组的任务是为 Internet 工作和发展提供技术及其他支持。它的任务之一是简化现在的标准和开发一些新的标准，并向 IESG 推荐标准。IETF 的目标是创建信息文档、创建协议细则，解决 Internet 与工程和标准制定有关的各种问题。

5）Internet 研究任务组(IRTF)

Internet 研究任务组是 ISOC 的执行机构。它致力于与 Internet 有关的长期项目的研究，主要在 Internet 协议、体系结构、应用程序及相关技术领域开展工作。

2. WWW 协会

WWW 协会是除 ISOC 之外的另一个国际性的 Internet 组织，其网址为 http://www.w3c.org。W3C 由美国麻省理工学院计算机科学实验室、欧洲国家信息与自动化学院和日本的 Keio University Shonan Fujisawa 联合组成。W3C 的主要任务在于确定和颁布有关 WWW 应用的标准。

3. 中国互联网络信息中心

中国互联网络信息中心(CNNIC)是 1997 年 6 月成立的非营利的 Internet 管理与服务机构，行使中国互联网络信息中心的职责。其网址为 http://www.cnnic.net.cn。

4. Internet 草案与 RFC

Internet 技术管理的核心是制定网络连接和应用的协议标准。在 Internet 上，任何一个用户都可以对 Internet 某一领域的问题提出自己的解决方案或规范，然后作为 Internet 草案(Internet Drafts)提交给 IETF，IETF 成员可对该草案进行讨论、测试和审查，最后，由 IESG 确定草案是否能够成为 Internet 的标准。RFC 文档是用于发布 Internet 标准和 Internet 其他正式出版物的一种网络文件或工作报告。RFC 文档初创于 1969 年，RFC 出版物由 RFC 编辑直接负责，并接受 IAB 的一般性指导。

9.2　应用层协议

在第 8 章中讲述了传输层为应用层提供了端到端的通信服务。但不同的网络应用进程之间，还需要不同的应用规则。因此，在传输层协议之上，还需要有应用层协议。这是因为，每个应用层协议都是为了解决某一类问题，而不同问题的解决又必须通过位于不同主机中的多个应用进程之间的通信和协同工作来完成。应用进程之间的这种通信必须遵守严格的规则。应用层的具体内容就是精确定义这些通信规则，具体来说，应用层协议应当做如下定义。

(1) 应用进程交换的报文类型，如请求报文和响应报文。

(2) 各种报文类型的语法，如报文中的各个字段及其详细描述。

(3) 字段的含义，即包含在字段中信息的含义。

(4) 进程何时、如何发送报文，以及对报文进行响应的规则。

互联网公共领域的标准应用的应用层协议是由 RFC 文档定义的，大家都可以使用。例如，万维网的应用层协议 HTTP(超文本传输协议)就是由 RFC 7230 定义的。如果浏览器开发者遵守 RFC 7230 标准，所开发出来的浏览器就能够访问任何遵守该标准的万维网服务器并获取相应的万维网页面。在互联网中还有很多其他应用的应用层协议不是公开的，而是专用的。例如，很多现有的 P2P 文件共享系统使用的就是专用应用层协议。

在这里，应用层协议与网络应用并不是同一个概念。应用层协议只是网络应用的一部分。例如，万维网应用是一种基于客户机/服务器体系结构的网络应用。万维网应用包含很多部件，有万维网浏览器、万维网服务器、万维网文档的格式标准，以及一个应用层协议。万维网的应用层协议是 HTTP，它定义了在万维网浏览器和万维网服务器之间传送的报文类型、格式和序列等规则。而万维网浏览器如何显示一个万维网页面，万维网服务器是用多线程还是用多进程来实现，则都不是 HTTP 所定义的内容。

应用层的许多协议都是基于客户机/服务器模式。即使是 P2P 对等通信方式，实质上也是一种特殊的客户机/服务器模式。这里再明确一下，客户端(Client)和服务器端(Server)都是指通信中所涉及的两个应用进程。客户机/服务器模式所描述的是进程之间服务和被服务的关系。这里最主要的特征就是：客户是服务请求方，服务器是服务提供方。

应用层是 TCP/IP 参考模型的最高层，其通过使用传输层所提供的服务，直接向用户提供服务，是 TCP/IP 网络与用户之间的界面或接口。该层由若干面向用户提供服务的应用协议和支持这些应用的支撑协议组成，基于这些协议，应用层向用户提供了众多的网络应用。

TCP/IP 应用层上的典型应用包括 Web 浏览、电子邮件、文件传输访问和远程登录等，与这些应用相关的协议包括超文本传输协议(HTTP)、简单邮件传输协议(SMTP)、文件传输协议(FTP)、简单文件传输协议(TFTP)和虚拟终端协议(Telnet)。

(1) HTTP：用来在浏览器和 WWW 服务器之间传送超文本的协议。

(2) SMTP：用于实现电子邮件传输的应用协议。

(3) FTP：用于实现文件传输服务的协议。通过 FTP 可以方便地连接到远程服务器上，可以进行查看、删除、移动、复制、更名远程服务器上的文件内容的操作，并能进行上传文件和下载文件等操作。

(4) TFTP：用于提供小而简单的文件传输服务。从某个意义上来说，TFTP是对FTP的一种补充，特别是在文件较小并且只有传输需求时该协议显得更加有效率。

(5) Telnet：实现虚拟或仿真终端的服务，允许用户把自己的计算机当作远程主机上的一个终端连接到远程计算机，并使用基于文本界面的命令控制和管理远程主机上的文件及其他资源。

为了使用户更加可靠、高效地访问网络应用服务，TCP/IP参考模型的应用层还提供了一些专门的应用支撑协议，如域名服务系统(DNS)、简单网络管理协议(SNMP)等。

(1) DNS：用于实现域名和IP地址之间的相互转换。

(2) SNMP：由于因特网结构复杂，拥有众多的操作者，因此需要好的工具进行网络管理，以确保网络运行的可靠性和可管理性。而SNMP提供了一种监控和管理计算机网络的有效方法，它已成为计算机网络管理的事实标准。

9.3 Internet的域名机制

IP地址为Internet提供了统一的编址方式，直接使用IP地址就可以访问Internet中的主机。但是一般用户很难记住地址编码，为此提出了域名这个概念，即服务器的地址用字符表示，每个字符都有一定的意义，并且书写有一定的规律，这样就容易理解并记忆了。于是当用户打开浏览器，在地址栏中输入域名后，就能看到IP地址所对应的页面。

可以看出域名和IP地址是一一对应的关系，网络上有一种叫作DNS服务器的计算机，它的作用就是自动把用户的域名"翻译"成相应的IP地址，并完成域名到IP地址的映射。当然DNS也承担着将IP地址解析成域名的任务，也就是反向解析。

9.3.1 DNS名称空间

DNS是域名服务系统(Domain Name System)的缩写，指在Internet中使用的分配名字和地址的机制。域名服务系统允许用户使用友好的名字而不是难以记忆的数字——IP地址来访问Internet上的主机。

域名解析就是将用户提出的名字变换成网络地址的方法和过程，从概念上讲，域名解析是一个自上而下的过程。

DNS域名采用分层结构，其域名结构的一般形式为"主机名.二级子域.子域.顶级域"结构形式。DNS实际上是一个分布式的数据库系统，某一台DNS并没有保存着所有主机信息的主机表，相反，这些信息都存放在许多分布式的域名服务器中，这些域名服务器组成一个自顶向下层次结构的系统。在整个系统的顶层是一个根域(Root Domain)，根域下面是众多一级和二级甚至三级域名。

9.3.2 域名解析过程

当客户机需要访问Internet上某一主机时，首先向本地DNS服务器查询对方的IP地

址，往往本地DNS服务器继续向另外一台DNS服务器查询，解析出访问主机的IP地址。这一过程称为“解析”。

域名解析是由一系列域名服务器来完成的。本质上，整个域名服务系统以一个大的分布式数据库的方式工作。大多数具有因特网连接的组织运行一台域名服务器。每台域名服务器包含连向其他域名服务器的信息，这些域名服务器连成一个大的域名数据库。域名服务器是运行在指定主机上的软件，能够完成从域名到IP地址的映射。分布式的主机信息数据库管理因特网上所有的主机域名和IP地址。不同的域名是分布式数据库的不同部分，每个域至少由一台域名服务器来管理。域名服务器保存有该域名空间的所有信息，并负责回答某个域名地址的查询要求。查询的结果可能有两个：若本地有该域名的地址则直接给出；若没有则给出其他相关域名服务器的地址。

根据域名层次来安排DNS服务器的层次，每台服务器作为域名体系中的一部分管辖者。一台根域名服务器只占这个层次体系的顶部，它是顶层域的管辖者。虽然根服务器并不包含所有可能的域名，但它知道如何查找相应请求的服务器。收到域名查询后，根域名服务器提供该域名所在的第一级域的域名服务器的地址，由此再提供该域名所在的二级域的域名服务器的地址，由此下去，直到找到答案。

域名解析中客户与服务器的交互过程如图9.2所示。

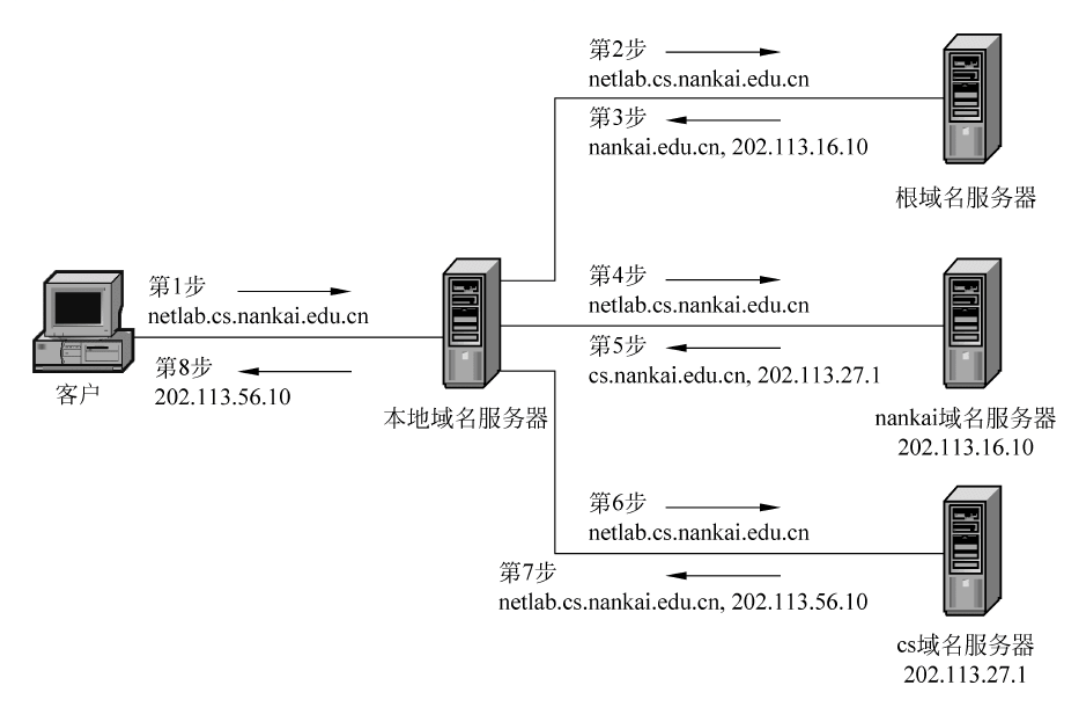


图9.2 域名解析中客户与服务器的交互过程

9.3.3 域名服务器的种类与查询方式

域名服务器主要有如下3类。

(1) 本地域名服务器：也被称为默认域名服务器。查询报文首先被送往该主机的本地域名服务器。当所要查询的主机也属于同一个本地ISP时，该本地域名服务器就能立即将所查询的主机名转换为IP地址，而不需要查询其他的域名服务器。

(2) 根域名服务器：目前在因特网上有十几台根域名服务器。当一台本地域名服务器不能满足某台主机的查询时，它就以 DNS 客户的身份向某一台根域名服务器查询。若根域名服务器有被查主机的信息，就发送 DNS 回答报文给本地域名服务器；若根域名服务器没有被查询主机的信息时，它就求助于某台保存有被查主机名字映射的授权域名服务器的 IP 地址。

(3) 授权域名服务器：每一台主机都必须在授权域名服务器处注册登记。通常，一台主机的授权域名服务器就是其本地 ISP 的一台域名服务器。实际上，为了更加可靠地工作，一台主机最好有至少两台授权域名服务器。许多域名服务器同时充当本地域名服务器和授权域名服务器。授权域名服务器总是可以将其管辖的主机名转换为该主机的 IP 地址。

DNS 允许两种类型的查询方式：递归查询和迭代查询。

(1) 递归查询。对于一个给定的主机名，若本地域名服务器不能满足查询要求，则求助于根域名服务器，而根域名服务器可能只知道中间域名服务器，而中间域名服务器又只知道授权域名服务器的 IP 地址。例如，A 要查询 B 的 IP 地址，当根域名服务器 D 收到 B 的主机名查询时，首先转发这个查询到中间域名服务器 E，这个域名服务器转发所有查询到授权域名服务器 F，它对所有 B 的主机名都是授权的。这个授权域名服务器将相应的映射发送到中间域名服务器，该服务器转发映射根域名服务器，而根域名服务器转发映射到本地域名服务器，该服务器又转发映射到发送请求的主机。该过程如图 9.3 所示。

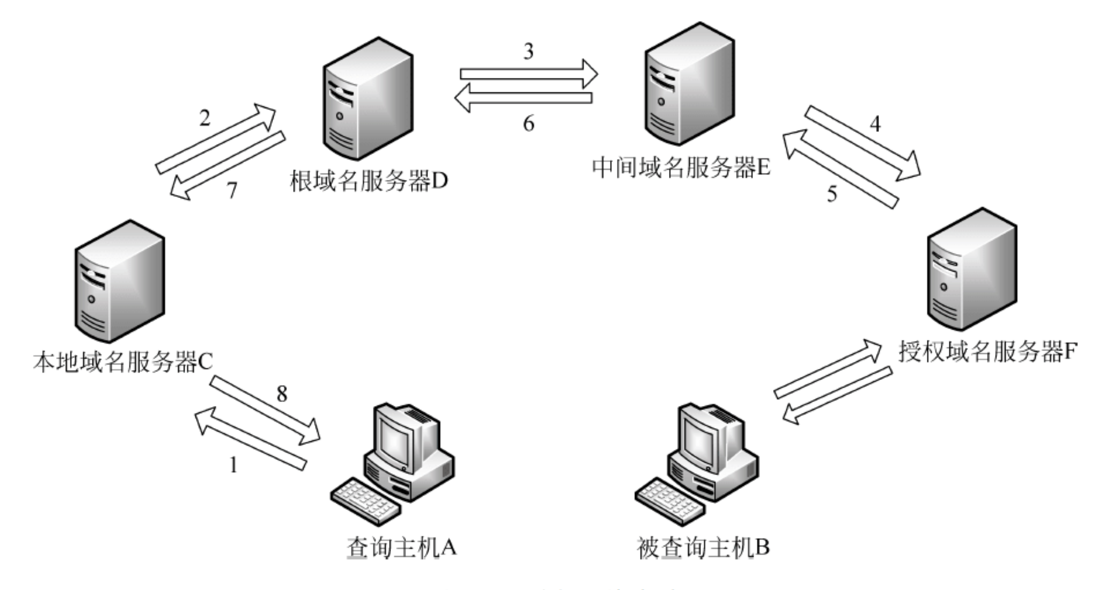


图 9.3 递归查询方式

(2) 迭代查询。当域名服务器 A 对 B 做了一个迭代查询后，假如 B 没有获得询问的映射，它立即送一个 DNS 到 A，该消息中包含了下一台域名服务器的 IP 地址 C，接下来，A 直接向 C 发送请求。当然，在实际的查询序列中，允许两者相结合进行查询。除了从本地域名服务器到根域名服务器的查询采用迭代查询外，其余查询采用递归查询，因为这样可以减轻根域名服务器的工作负担，如图 9.4 所示。

为了提高名称解析效率，将一次解析的结果保存在服务器的缓存之中，下次请求查询相同目标时，服务器端将直接从缓存中取出查询结果返回给客户端即可。每台服务器都保留

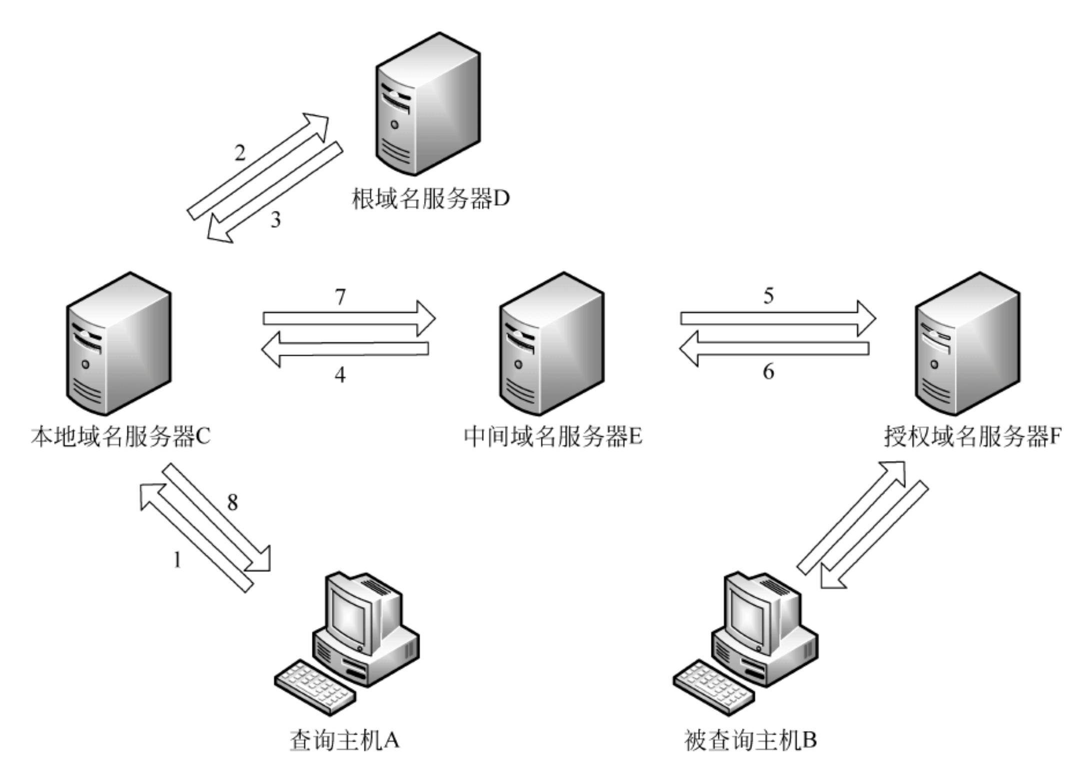


图 9.4　递归查询方式与迭代查询方式的结合

一个域名缓存，每当查找一个新的域名时，服务器将它置于缓存中；主机在启动时从本地域名服务器下载域名和地址的全部数据库，维护存放自己最近使用的域名的高速缓存，并且只在从缓存中找不到域名时才使用域名服务器。

域名缓存至少可以带来如下几方面的好处。

(1) 域名解析的高效性：本地的一个客户端在解析过某个域名后，其他的客户端对该域名的解析就可以在本地的域名解析服务器缓存中快速命中，极大地提高域名解析的效率。

(2) 域名解析的分布性：通过地理上分散的域名解析服务器缓存的解析处理，将最终递归或迭代到域名管理服务器的域名解析处理结果分布到当地(或就近)的域名解析服务器进行处理，这极大地缓解了该域名管理服务器的解析处理压力。

(3) 域名解析的可靠性：通过分布的解析服务器的协调处理，提高了整个域名解析系统的可靠性，局部的域名服务器故障不会导致整个域名服务系统的瘫痪。

9.4　Web 服务

万维网是因特网上发展最快同时又使用最多的一项服务，它可以提供包括文本、图形、声音和视频等在内的多媒体信息的浏览。

万维网起源于 1989 年欧洲粒子物理研究室(CERN)。其目的是收集时刻变化的报告、蓝图、绘制图、照片和其他文献。链接文档的万维网 Web 的最初计划是由 CERN 的物理学家 Tim Berners-Lee 于 1989 年 3 月提出的，第一个原型(基于文本的)于 18 个月后运行。1991 年 12 月在得克萨斯州的 San Antonio 91 超文本会议上进行了一次公开演示，次年继

续发展，并于1993年2月，在第一个图形界面Mosaic的发布时达到了其发展的高峰，现在WWW已经成为因特网上不可缺少的主流应用。

9.4.1 Web的基本概念

WWW由遍布在因特网中的被称为WWW服务器（又称为Web服务器）的计算机组成，是一个容纳各种类型信息的集合。从用户的角度看，Web由庞大的、世界范围的文档集合而成，简称为页面（Page）。页面具有严格的格式，页面是用超文本标记语言（Hyper Text Markup Language，HTML）写成的，存放在Web服务器上。每一页面可以包含到世界上任何地方的其他相关页面的超链接（Hyperlink），这种能够指向其他页面的页被称为超文本（Hypertext）。

用户使用浏览器总是从访问某个主页（Homepage）开始的。由于页面中可能包含了超链接，所以用户可以跟随超链接到它所指向的其他页面，并且这一过程可以被无限制地重复。通过这种方法可浏览到大量的相互链接的信息。下面来介绍一下WWW中常用的一些术语。

1. 超文本标记语言

超文本标记语言（HTML）是ISO标准8879-标准通用标记语言（Standard Generalized Markup Language，SGML）在万维网上的应用。所谓标记语言就是格式化的语言，存在于WWW服务上的页，就是由HTML描述的。它使用一些约定的标记对WWW上各种信息（包括文字、声音、图形、图像、视频等）、格式以及超链接进行描述。当用户浏览WWW上的信息时，浏览器会自动解释这些标记的含义，并将其显示为用户在屏幕上所看到的网页。

一个Html文本包括文件头（Head）、文件（Body）主体两部分。其结构如下所示。

```
<Html>
    <Head>
    </Head>
        <Body>
        …
        …
        </Body>
</Html>
```

其中，<Html>表示页开始，</Html>表示页结束，它们是成对使用的；<Head>表示头开始，</Head>表示头结束；<Body>表示主体开始，</Body>表示主体结束，它们之间的内容才会在浏览器的正文中显示出来。Html的标识符有很多，有兴趣的同学可以查看有关网页制作方法的书籍。

2. 超文本传输协议

超文本传输协议（Hypertext Transfer Protocol，HTTP）是用来在浏览器和WWW服务器之间传送超文本的协议，它可以使浏览器更加高效，使网络传输减少。HTTP是一种面向对象的协议，它由两部分组成：从浏览器到服务器的请求集和从服务器到浏览器的应答集。为了保证WWW客户机与WWW服务器之间通信不会产生二义性，HTTP精确定义了请求报文和响应报文的格式。HTTP会话过程包括连接、请求、应答和关闭4个步骤，

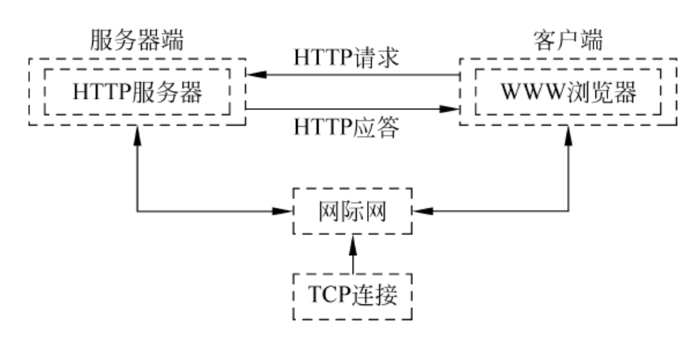


图 9.5 HTTP 会话过程

如图 9.5 所示。

HTTP 从层次的角度看，是面向事务的应用层协议。它是万维网上能够可靠地交换文件(包括文本、声音、图像等各种多媒体文件)的重要基础。该协议不仅保证计算机正确快速地传输超文本文档，还准确了传输文档中位置等。所以在浏览器中看到的网页地址都是以"http：//"开头的。

由于 HTTP 是基于客户机/服务器模式，所以客户机与服务器建立连接后，发送一个请求给服务器，请求方式的格式为：统一资源标识符(URL)、协议版本号，后面是 MIME 信息，包括请求修饰符、客户机信息和可能的内容。服务器接到请求后，给予相应的响应信息，其格式为一个状态行，包括信息的协议版本号、一个成功或错误的代码，后面是 MIME 信息，包括服务器信息、实体信息和可能的内容。归结起来，HTTP 有如下 6 个重要特点。

(1) 采用客户机/服务器模型：HTTP 的设计支持客户机/服务器通信，注重超文本数据的传输。单台服务器可以为世界范围内众多的客户提供信息服务。

(2) 简易性：HTTP 使得 WWW 服务器能够高速地处理大量请求，客户通过发送请求方式和 URL 等规格化信息就能使用服务，与 FTP 等协议相比，HTTP 速度快、开销小。

(3) 灵活性与扩展性：HTTP 允许传送任意类型的数据，在 HTTP 的信息包中，通过内容/类型标识可以定义传输的数据类型，不同的数据贴上不同的标签，就可指明操作方法。随着新的数据格式涌现，HTTP 只需公布新的标识就可以为这些数据传送提供服务。

(4) 无连接性：HTTP 就好像是寄信，服务器收到一封申请信，马上答复一封信，每一次在服务器一方都是独立的，不需要在请求的间隔中浪费时间。

(5) HTTP 的无状态性表现在两方面：一方面协议不记忆事务，为后续事务所需的信息必须在协议之外完成，从而每一次都需要传递完全的信息来说明服务，有些必要信息可能大量重复传送；另一方面，HTTP 无须每次保留维护状态表，可以加快处理速度。

(6) HTTP 在客户方提出请求时，可以指明能够接受的响应类型，从而在服务器一方可以用最恰当的方式把信息组合起来送交客户方。

3. 统一资源定位器(URL)

WWW 是以页面的形式来组织信息的。那么怎样来识别不同的页面，怎样才能知道页面在哪个位置，以及如何访问页面呢？为了解决这个问题，WWW 采用了统一资源定位器(Uniform Resource Locator，URL)的方法。

URL 是在因特网上唯一确定资源位置的方法，其基本格式如下：

协议：//主机域名/资源文件名

其中，“协议(Protocol)”用来指明资源类型，除了 WWW 用的 HTTP 协议之外，还可以是 FTP、Telnet 等协议；“主机域名”表示资源所在机器的 DNS 名字；“资源文件名”用以提出资源在所在机器上的位置，包含路径和文件名，通常为“目录名/子目录名/文件名”，也可以不含有路径。例如，西安交通大学的 WWW 主页的 URL 就表示为“http://www.xjtu.edu.cn/index.php”。

在输入 URL 时，协议和主机域名不分字母的大小写，但目录和文件名则可能区分字母的大小写。这是因为大多数服务器安装了 UNIX 操作系统，而 UNIX 的文件系统是区分文件名的大小写的。

9.4.2 WWW 服务的实现过程

WWW 以客户端/服务器端(Client/Server)的模式进行工作。运行 WWW 服务器程序并提供 WWW 服务的机器被称为 WWW 服务器；在客户端，用户通过一个被称为浏览器(Browser)的交互式程序来获得 WWW 服务。常用到的浏览器有 Safari、Firefox、Chrome 和 Internet Explorer 等。

在服务器端，对于每个 WWW 服务器站点，都有一个关于 TCP 的 80 端口的监听(注：80 为 HTTP 默认的 TCP 端口)，看是否有从客户端(通常是浏览器)过来的连接。在客户端，当浏览器在其地址栏里输入一个 URL 或者单击 Web 页上的一个超链接时，Web 浏览器就要通过解析器对域名进行解析以获得相应的 IP 地址。然后，以该 IP 地址为目标地址，以 HTTP 所对应的 TCP 端口为源端口与服务器端建立一个 TCP 连接。连接建立之后，客户端的浏览器使用 HTTP 中的 GET 功能向 WWW 服务器发出指定的 WWW 页面请求，服务器端收到该请求后将根据客户端所要求的路径和文件名使用 HTTP 中的 PUT 功能将相应的 HTML 文档回送到客户端，如果客户端没有指明相应的文件名，则由服务器端返回一个默认的 HTML 页面。页面传送完毕后，中止相应的 TCP 连接。

下面以一个具体的例子来说明 Web 服务的实现过程。假设有用户要访问温州大学主页 http://www.wzu.edu.cn/index.php，则浏览器与服务器的信息交互过程如下。

(1) 浏览器确定 URL。

(2) 浏览器向 DNS 获取 Web 服务器 www.wzu.edu.cn 的 IP 地址。

(3) DNS 服务器以相应的 IP 地址 218.75.16.107 应答。

(4) 浏览器和 IP 地址为 218.75.16.107 的主机的 80 端口建立一条 TCP 连接。

(5) 浏览器执行 HTTP，发送 GET“/index.php”命令，请求读取该文件。

(6) www.wzu.edu.cn 服务器返回“/index.php”文件到客户端。

(7) 释放 TCP 连接。

(8) 浏览器显示“/index.php”中的所有正文和图像。

自 WWW 服务问世以来，其已取代电子邮件服务成为因特网上最为广泛的服务。除了普通的页面浏览外，WWW 服务中的浏览器/服务器(Browse/Server，B/S)模式还取代了传统的 C/S 模式，被广泛用于网络数据库应用开发中。

9.5 FTP 服务

FTP(File Transfer Protocol)是文件传输协议的简称。FTP 的主要作用就是让用户连接上一台远程计算机(这些计算机上运行着 FTP 服务器程序),查看远程计算机有哪些文件,然后把文件从远程计算机上复制到本地计算机,或把本地计算机的文件送到远程计算机上。

当启动 FTP 从远程计算机复制文件时,实际上启动了两个程序:一个本地机上的 FTP 客户程序:它向 FTP 服务器提出复制文件的请求。另一个是启动在远程计算机上的 FTP 服务器程序,它响应用户的请求,把用户指定的文件传送到计算机中。FTP 采用客户机/服务器模式,用户端要在自己的本地计算机上安装 FTP 客户程序。FTP 客户程序有字符界面和图形界面两种。字符界面的 FTP 的命令复杂、繁多。图形界面的 FTP 客户程序,操作上简洁、方便。

FTP 采用两个默认端口号:20 和 21。其中,20 号端口用于数据传输,21 号端口用于控制信息的传输。控制信息和数据信息能够同时传输,是 FTP 的特殊之处。FTP 的文件传输处理过程如图 9.6 所示。

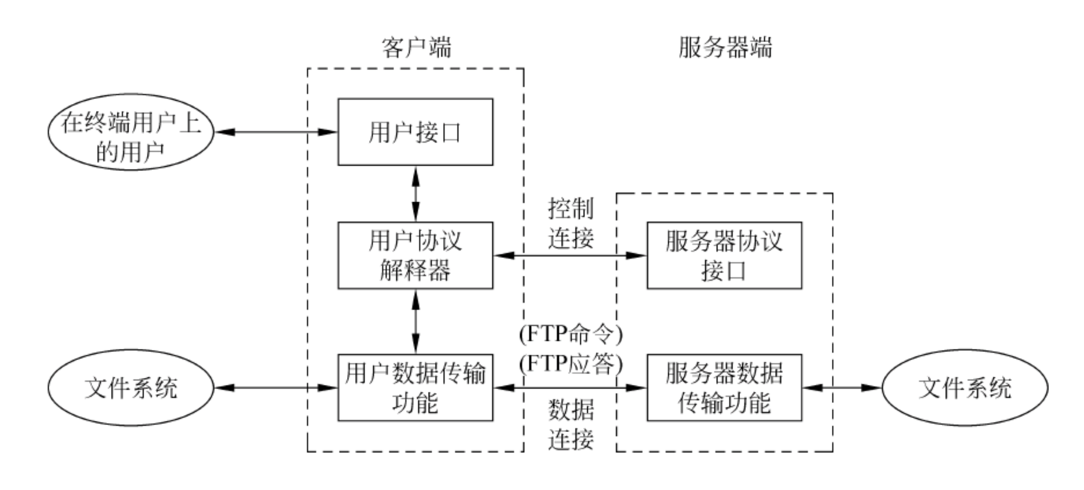


图 9.6 文件传输的处理过程

简单地说,支持 FTP 的服务器就是 FTP 服务器,下面介绍一下什么是 FTP(文件传输协议)。

一般来说,用户联网的首要目的就是实现信息共享,文件传输是信息共享非常重要的一个内容。Internet 上早期实现传输文件,并不是一件容易的事,我们知道 Internet 是一个非常复杂的计算机环境,有 PC、工作站、大型机,据统计,连接在 Internet 上的计算机已有上千万台,而这些计算机可能运行不同的操作系统,有运行 UNIX 的服务器,也有运行 DOS、Windows 的 PC 和运行 Mac OS 的苹果机等,而各种操作系统之间的文件交流问题,需要建立一个统一的文件传输协议,这就是所谓的 FTP。基于不同的操作系统有不同的 FTP 应用程序,而所有这些应用程序都遵守同一种协议,这样用户就可以把自己的文件传送给别人,或者从其他的用户环境中获得文件。

与大多数 Internet 服务一样，FTP 也是一个客户机/服务器系统。用户通过一个支持 FTP 的客户机程序，连接到远程主机上的 FTP 服务器程序。用户通过客户机程序向服务器程序发出命令，服务器程序执行用户所发出的命令，并将执行的结果返回到客户机。比如说，用户发出一条命令，要求服务器向用户传送某一个文件的一份副本，服务器会响应这条命令，将指定文件送至用户的机器上。客户机程序代表用户接收到这个文件，将其存放在用户目录中。

在 FTP 的使用当中，用户经常遇到两个概念："下载"(Download)和"上传"(Upload)。"下载"文件就是从远程主机复制文件至自己的计算机上；"上传"文件就是将文件从自己的计算机中复制至远程主机上。用 Internet 语言来说，用户可通过客户机程序向(从)远程主机上传(下载)文件。

使用 FTP 时必须首先登录，在远程主机上获得相应的权限以后，方可上传或下载文件。也就是说，要想同哪一台计算机传送文件，就必须具有哪一台计算机的适当授权。换言之，除非有用户 ID 和口令，否则便无法传送文件。这种情况违背了 Internet 的开放性，Internet 上的 FTP 主机何止千万，不可能要求每个用户在每一台主机上都拥有账号。匿名 FTP 就是为解决这个问题而产生的。

匿名 FTP 是这样一种机制，用户可通过它连接到远程主机上，并从其下载文件，而无须成为其注册用户。系统管理员建立了一个特殊的用户 ID，名为 anonymous，Internet 上的任何人在任何地方都可使用该用户 ID。通过 FTP 程序连接匿名 FTP 主机的方式同连接普通 FTP 主机的方式相似，只是在要求提供用户标识 ID 时必须输入 anonymous，该用户 ID 的口令可以是任意的字符串。习惯上，用自己的 E-mail 地址作为口令，使系统维护程序能够记录下来谁在存取这些文件。值得注意的是，匿名 FTP 不适用于所有 Internet 主机，它只适用于那些提供了这项服务的主机。

当远程主机提供匿名 FTP 服务时，会指定某些目录向公众开放，允许匿名存取。系统中的其余目录则处于隐匿状态。作为一种安全措施，大多数匿名 FTP 主机允许用户从其下载文件，而不允许用户上传文件，也就是说，用户可将匿名 FTP 主机上的所有文件全部复制到自己的机器上，但不能将自己机器上的任何一个文件复制至匿名 FTP 主机上。即使有些匿名 FTP 主机确实允许用户上传文件，用户也只能将文件上传至某一指定上传目录中。随后，系统管理员会去检查这些文件，会将这些文件移至另一个公共下载目录中，供其他用户下载，利用这种方式，远程主机的用户得到了保护，避免了有人上传有问题的文件，如带病毒的文件。

作为一个 Internet 用户，可通过 FTP 在任何两台 Internet 主机之间复制文件。但是，实际上大多数人只有一个 Internet 账户，FTP 主要用于下载公共文件，例如共享软件、各公司技术支持文件等。Internet 上有成千上万台匿名 FTP 主机，这些主机上存放着数不清的文件，供用户免费复制。实际上，几乎所有类型的信息，所有类型的计算机程序都可以在 Internet 上找到。这是 Internet 吸引我们的重要原因之一。

匿名 FTP 使用户有机会存取到世界上最大的信息库，这个信息库是日积月累起来的，并且还在不断增长，永不关闭，涉及几乎所有主题。匿名 FTP 是 Internet 网上发布软件的常用方法。Internet 之所以能延续到今天，是因为人们使用通过标准协议提供标准服务的程序。像这样的程序，有许多就是通过匿名 FTP 发布的，任何人都可以存取它们。

Internet 中有数目巨大的匿名 FTP 主机以及更多的文件，那么到底怎样才能知道某一特定文件位于哪个匿名 FTP 主机上的哪个目录中呢？Archie 将自动在 FTP 主机中进行搜索，构造一个包含全部文件目录信息的数据库，使你可以直接找到所需文件的位置信息。

9.6　E-mail 服务

电子邮件(Electronic Mail，E-mail)是因特网上最受欢迎也颇为广泛的应用之一。电子邮件服务(E-mail)是一种通过计算机网络与其他用户进行联系的快速、简便、高效、廉价的现代化通信手段。电子邮件之所以受到广大用户的喜爱，是因为与传统通信方式相比，其具有以下明显的优点。

(1) 成本低。与传统的邮件系统相比，电子邮件费用很低。传统国内特快需 20 元人民币，国际快递则更贵；而通过电子邮件将信件发送至国外，可能只需付几分钱的上网费。

(2) 速度快。电子邮件一般只需几秒钟就可以到达目的地，远比人工邮件传递速度要迅速，而且比较可靠。

(3) 安全与可靠性高。使用电子邮件不必担心损坏，传统的邮件在投递过程中，有可能信件被损坏，而使用电子邮件则不必担心这一点。

(4) 可达到范围广。电子邮件可以到达因特网可达的任何地方，并且可以实现一对多的邮件传送，即可以一次同时向多人发出多个内容相同的邮件。

(5) 内容表达形式多样。电子邮件可以将文字、图像、语音等多种类型的信息集成在一个邮件中传送，因此它成为多媒体信息传送的重要手段。

那么电子邮件是如何通过网络被发送和接收出去的呢？首先电子邮件要有自己规范的格式，就好比使用普通的邮件系统要遵循标准的邮件格式一样。

9.6.1　电子邮件格式

电子信箱实际上就是在 Internet 服务商(Internet Service Provider，ISP)的电子邮件服务器上为用户开辟出一块专用的磁盘空间，用来存放用户的电子邮件文件。每个电子信箱都有一个地址，称为电子信箱地址(E-mail Address)。电子信箱地址的格式是固定的，并且在全球范围内是唯一的。

E-mail 地址的标准格式为：＜用户名＞@主机域名。其中，用户名是指用户在某台邮件服务器上注册的用户标识，相当于是他的一个私人邮箱，用户名通常由用户自行选定，但在同一个邮件服务器上必须是唯一的；@为分隔符，一般把它读为英文的“at”；主机域名是指信箱所在的邮件服务器的域名，域名可以由几部分组成，每一部分称为一个子域，各子域之间用圆点“.”隔开，每个子域都会告诉用户一些有关这台邮件服务器的信息。例如 zcr@mail.xjtu.edu.cn，表示在西安交通大学的邮件服务器上的用户名为 zcr 的用户信箱，即这台计算机在中国(cn)，并隶属于教育机构(edu)下的西安交通大学(xjtu)，机器名(mail)。在@符号的左边是用户名(zcr)。

电子邮件的格式有信封和内容两大部分，即邮件头(Header)和邮件主体(Body)。邮件头包括收信人的 E-mail 地址、发信人的 E-mail 地址、发送日期、标题和发送优先级等；其中，前两项是必选的。邮件主体才是发件人和收件人要处理的内容，早期的电子邮件系统只能传递文本信息，而通过使用多用途因特网邮件扩展协议(Multipurpose Internet Mail Extensions，MIME)，现在还可以发送语音、图像和视频等信息。对于 E-mail 主体不存在格式上的统一要求，但对信封即邮件头有严格的格式要求，尤其是 E-mail 地址。

除了标准的电子邮件格式外，电子邮件的发送与接收还要依托由用户代理、邮件服务器和邮件协议组成的电子邮件系统。图 9.7 给出了电子邮件系统的简单示意图。

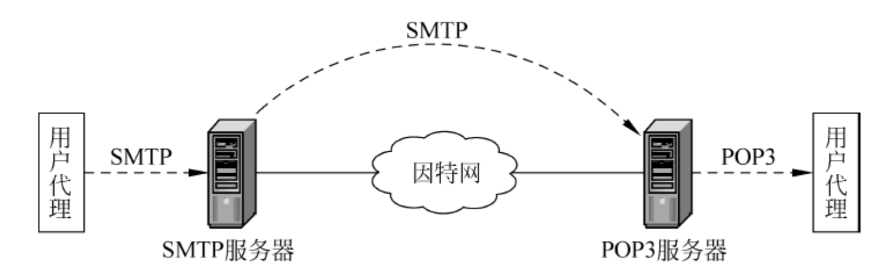


图 9.7 电子邮件系统的构成

其中，用户代理运行在客户机上的一个本地程序，它提供命令行方式、菜单方式或图形方式的界面来与电子邮件系统交互，允许人们读取和发送电子邮件，如 Outlook Express 或 Hotmail 等。邮件服务器包括邮件发送服务器和邮件接收服务器。顾名思义，所谓邮件发送服务器，是指为用户提供邮件发送功能的邮件服务器，如图 9.7 中的 SMTP 服务器；而邮件接收服务器是指为用户提供邮件接收功能的邮件服务器，如图 9.7 中的 POP3 服务器。用户在发送邮件时，要使用邮件发送协议，常见的邮件发送协议有简单邮件传输协议(Simple Mail Transfer Protocol，SMTP)和 MIME，前者只能传输文本信息，后者则可以传输包括文本、声音、图像等在内的多媒体信息。

9.6.2 电子邮件原理

电子邮件在发送和接收过程中，还要遵循一些基本协议和标准，如 SMTP、POP3 等，这些协议和标准保证电子邮件在各种不同系统之间进行传输。

SMTP 即简单邮件传输协议，它是一组用于由源地址到目的地址传送邮件的规则，由它来控制信件的中转方式。SMTP 属于 TCP/IP 协议簇，它帮助每台计算机在发送或中转信件时找到下一个目的地。通过 SMTP 所指定的服务器，就可以把 E-mail 寄到收信人的服务器上了，整个过程只要几分钟。SMTP 服务器则是遵循 SMTP 的发送邮件服务器，用来发送电子邮件。

POP3(Post Office Protocol 3)即邮局协议的第 3 个版本，它规定怎样将个人计算机连接到 Internet 的邮件服务器和下载电子邮件的电子协议。它是因特网电子邮件的第一个离线协议标准，POP3 允许用户从服务器上把邮件存储到本地主机(即自己的计算机)上，同时删除保存在邮件服务器上的邮件，而 POP3 服务器则是遵循 POP3 的接收邮件服务器，用来接收电子邮件的。通常，SMTP 使用 TCP 的 25 号端口，而 POP3 则使用 TCP 的 110 号

端口。

图 9.8 给出了一个电子邮件发送和接收的具体实例。假定用户 XXX 使用“XXX@sina. com”作为发信人地址向用户 YYY 发送一个文本格式的电子邮件，该发信人地址所指向的邮件发送服务器为 smtp. sina. com. cn，收信人的 E-mail 地址为“YYY@163. net”。

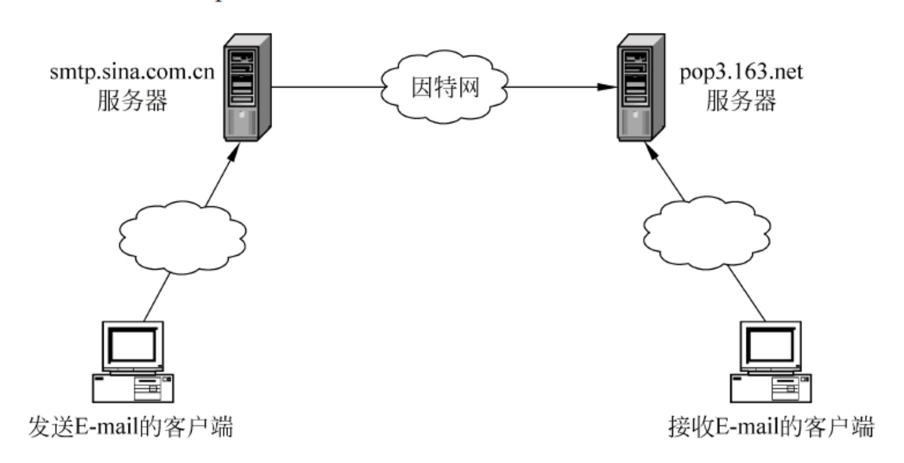


图 9.8 电子邮件发送和接收实例

首先，用户 XXX 在自己的机器上使用独立式的文本编辑器、字处理程序或是用户代理内部的文本编辑器来撰写邮件正文；其次，使用电子邮件用户代理程序如 Outlook Express 完成标准邮件格式的创建，即选择创建新邮件图标，填写收件人地址、主题、邮件的正文、邮件的附件等。

一旦用户邮件发送图标之后，则“用户代理程序”将用户的邮件传给负责邮件传输的程序，由其在 XXX 所用的主机和名为 smtp. sina. com. cn 的发送服务器之间建立一个关于 SMTP 的连接，并通过该连接将邮件发送至服务器 smtp. sina. com. cn。

发送方服务器 smtp. sina. com. cn 在获得用户 XXX 所发送的邮件后，根据邮件接收者的地址，在发送服务器与用户 YYY 的接收邮件服务器之间建立一个 SMTP 的连接，并通过该连接将邮件发送至用户 YYY 的接收服务器。

接收方邮件服务器 smtp. 163. net 接收到邮件后，根据邮件接收者的用户名将邮件放到用户的邮箱中。在电子邮件系统中，为每个用户分配一个邮箱（用户邮箱）。例如，在基于 UNIX 的邮件服务系统中，用户邮箱位于/usr/spool/mail/目录下，邮箱标识一般与用户标识相同。

当邮件到达邮件接收服务器后，用户随时都可以接收邮件。当用户 YYY 需要查看自己的邮箱并接收邮件时，其首先要在自己的机器与邮件接收服务器 POP3. 163. net 之间建立一条关于 POP3 的连接，该连接也是通过系统提供的“用户代理程序”进行。连接建立之后，用户就可以从自己的邮箱中“取出”邮件进行阅读、处理、转发或回复邮件等操作。

从上面的例子可以看出，电子邮件的“发送—传递—接收”是异步的，邮件的发送时并不要求接收者在线（Online），邮件可存放在接收用户的邮箱中，接收者随时可以上网接收。

一般来说，电子邮件系统支持以下功能。

（1）撰写（Composition）：提供一个非常方便的编辑信件的环境，来创建消息和回答的过程。

（2）传输（Transfer）：将信件从发送方传输到接收方。

（3）报告（Reporting）：告诉发送方信件发送的情况。

（4）显示（Displaying）：到接收方以后，应显示信件内容。有时需要进行转换或者需要激活浏览器。

（5）处理（Disposition）：当接收方阅读信件后，需要将该邮件进行处理。例如，丢弃或者保存等。

课后习题

1．应用层有哪些主要协议？每个协议的主要作用是什么？
2．Internet 的域名解析有几种方式？分别解释。
3．域名解析过程中的域名缓存有哪些优点？
4．解释 HTTP 协议的会话过程。
5．什么是 URL？其基本格式如何表示？
6．发送电子邮件的格式是什么？解释电子邮件系统的构成。

第10章 网络管理与网络安全

学习目的

本章系统地学习网络管理技术以及网络安全的基本概念、密码加密和数字证书的基本概念、防火墙、入侵检测、VPN技术以及网络病毒防治的基本方法与技术。通过本章的学习，初步建立网络管理的技术要点以及网络安全的基本概念，掌握密码加密、防火墙、网络入侵检测、VPN技术以及网络病毒防治的技术要点。

学习要求

了解：网络安全的重要性。

了解：网络管理的基本概念与方法。

掌握：密码体制的基本概念及应用。

掌握：防火墙的基本概念。

掌握：网络入侵检测的基本概念与方法。

了解：网络病毒防治的基本方法与技术。

10.1 网络管理的基本概念

网络管理是指网络管理员通过网络管理程序对网络上的资源进行集中化管理的操作，包括配置管理、性能和记账管理、问题管理、操作管理和变化管理等。一台设备所支持的管理程度反映了该设备的可管理性及可操作性。

计算机网络的管理是与Internet同步发展的。随着计算机技术的发展，网络规模逐渐增大，复杂性越来越高，在这种环境下，资源分布程度和共享程度大大提高，任何微小的故障都可能导致用户应用的失败。如何及早发现并排除潜在的故障隐患、有效地管理好网络是网络建设者、服务提供者共同关心的问题。

1. 网络管理的目标

网络管理是提高网络安全性、可靠性的技术保证，也是提高网络效益的一种方式，其目的最终要实现的是最大限度地增加网络的可用时间，提高网络设备的利用率、网络性能、服务质量和安全性，简化多厂商混合网络环境下的管理和控制网络运行的成本，减少或消除网络瓶颈，适应新技术，使网络更容易、安全，并对网络提供长期规划。

2. 网络管理的任务

（1）状态监测。通过状态监测，可以获得分析网络各种性能的原始数据。

（2）数据收集。要了解网络的状态，还需要将分散监测到的有用数据收集到一起。

（3）状态分析。利用各种模型，根据收集到的监测数据对网络的状态进行分析、判断。

（4）状态控制。根据状态分析的结果对网络采取控制措施。

3. 网络管理的功能

1）配置管理

配置管理是指从网络中获取信息并根据这些信息来对设备进行配置管理，它是网络管理最基本的功能。在一个计算机网络中所用到的网络互联设备通常来自不同的厂商，而各个厂商设备之间需要进行设备的参数、状态信息等内容的相互交换工作；同时网络是不断发展变化的，如网络系统要随着用户数量、设备更新来调整网络的配置。因此，需要有网络配置管理功能来支持这种调整和改变，保证网络系统的正常运行。包括以下几项功能。

（1）更改系统设置。

（2）获取系统重要变化信息。

（3）获取系统状态信息。

（4）初始化或关闭管理对象。

2）故障管理

故障管理是指通过检测、隔离、修复网络故障等方法，使网络恢复正常运行状态。与单台计算机不同，在计算机网络中，当某个组成部分发生故障时，往往不能及时准确地确定故障发生的位置。因此，需要一个具备故障管理功能的工具，来科学地管理网络出现的所有故障，并记录故障产生位置及其相关信息，以达到快速、准确解决网络故障，保证网络提供连续可靠服务的目的。包括以下几项功能。

（1）检测被管对象的差错，或接收差错报告。

（2）创建和维护差错日志库，并对其进行分析。

（3）进行诊断，以跟踪和识别差错。

（4）通过恢复措施，恢复正确的网络服务。

（5）网络实时备份。

3）计费管理

计费管理是指在计算机网络系统资源为有偿使用的情况下，对用户使用网络信息资源的情况进行记录和统计，以控制和监测网络操作的费用与代价。即使是在非商业化用途的计算机网络中，仍需要计费管理来对网络资源的利用率及与其相关的信息进行统计，以便于管理员实时掌握网络的运行状态。包括以下几项功能。

（1）记录和统计网络资源利用率。

（2）设置计费标准。

（3）联机收集计费数据。

（4）计算用户应支付的费用。

（5）账单管理。

4）安全管理

安全管理是指通过采用信息安全措施以保证计算机网络系统资源不被非法使用、防止未经授权的访问和保护网络资源的完整性。一般来讲，安全管理具有的功能包括：风险分析、访问权限控制、安全服务、警告、日志和报告等。包括以下几项功能。

（1）安全措施相关的信息分发。

（2）安全相关的事件通知。

（3）安全相关的设施建设、控制和删除。

（4）涉及安全服务的网络操作事件的记录、维护和查阅等日志管理工作。

5）性能管理

性能管理是指通过分析和控制整个网络的数据交换，保证网络能够提供持续可靠的服务，使网络达到最好的运营效率。性能管理评测系统资源的运行状况和通信效率等系统性能，并收集分析当前状况的网络信息，维护和分析系统性能日志，保证在最小网络消耗和网络时延下，提供最大的可靠且连续的通信能力。包括以下几项功能。

（1）从被管对象中收集与网络性能有关的数据。

（2）对收集到的数据进行统计分析，并对历史记录进行维护。

（3）分析数据，以检测性能故障，生成报告。

（4）预测网络的长期趋势。

（5）改进网络的操作模式。

10.2 简单网络管理协议

简单网络管理协议（Simple Network Manage Protocol，SNMP）首先是由 Internet 工程任务组（Internet Engineering Task Force，IETF）的研究小组为了解决 Internet 上的路由器管理问题而提出的。它可以在 IP、IPX、AppleTalk、OSI 以及其他用到的传输协议上使用。SNMP 是一系列网络管理规范的集合，包括协议本身、数据结构的定义和一些相关概念。

SNMP 是很早提出的网络管理协议之一，一推出就得到了广泛的应用和支持，特别是很快得到了数百家厂商的支持，其中包括 IBM、HP、SUN 等大公司和厂商。目前，SNMP 已成为网络管理领域中常用的工业标准，并被广泛支持和应用，大多数网络管理系统和平台是基于 SNMP 的。

现在 SNMP 是第三个版本的协议，其功能已经得到了加强和改进。SNMP 的体系结构是围绕着 4 个概念和目标进行设计的，即保证管理代理的软件成本尽可能低；最大限度地保持远程管理的功能，以便充分利用 Internet 的网络资源；体系结构必须有扩充的余地；保持 SNMP 的独立性，不依赖于具体的计算机、网关和网络传输协议。

10.2.1 SNMP 模型

SNMP 作为一种网络管理协议，使网络设备彼此可以交换管理信息，使网络管理员能够了解网络的性能、定位和解决网络故障，进行网络规划。SNMP 的体系结构分为 SNMP 管理（SNMP Manager）和 SNMP 代理（SNMP Agent），每个支持 SNMP 的网络设备中都包括一个代理，此代理随时记录网络设备的各种情况，网络管理程序再通过 SNMP 通信协议查询或修改代理所记录的信息。

SNMP 的网络管理模型由 3 个重要元素组成，如图 10.1 所示。

1. 管理信息库

被管理的网络设备都具有若干个变量来描述状态，这些变量称为对象。例如，路由器有

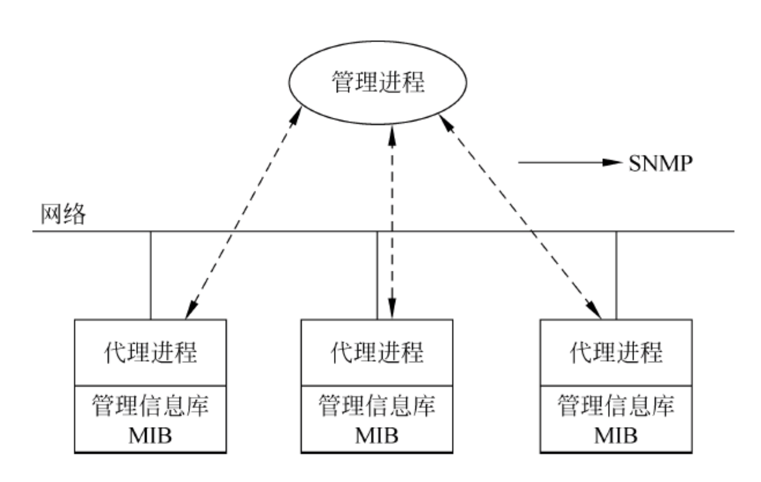


图 10.1　SNMP 的网络管理模型

每个端口发送和接收到的报文数量、丢弃的报文数量、IP 路由表等对象。网络中所有可供存取的对象存放在一个管理信息库(Management Information Base,MIB)中。

Internet 的 MIB 标准把管理对象分为 9 组,每组对象由各种变量描述。这 9 个组是 system、interface、at、ip、icmp、tcp、udp、egp 和 snmp,如表 10.1 所示,现在 at 组和 egp 组已不再使用。其余大部分组的定义都移到各自的文档中。

表 10.1　最初节点 MIB 管理的信息类型

类　　型	标　号	信 息 类 型
system	1	主机或路由器的操作系统
interface	2	各种网络接口及它们的测定通信量
address translation(at)	3	地址转换(如 ARP 映射)
ip	4	Internet 软件(IP 分组统计)
icmp	5	ICMP 软件(已收到 ICMP 消息的统计)
tcp	6	TCP 软件(算法、参数和统计)
udp	7	UDP 软件(UDP 通信量统计)
egp	8	EGP 软件(外部网关协议通信量统计)
snmp	9	SNMP 软件

MIB 中对象的变量类型大体可分为两种:简单变量和表格。简单变量包括整型变量、字符串变量等,也包括一些类似于 C 语言"结构"的数据集合,表格相当于一维数组。值得注意的是,MIB 只给出每个变量的逻辑定义,每个被管的网络设备所使用的内部数据结构可能与 MIB 的定义不同。当代理进程收到查询请求时,先要把 SNMP 的 MIB 变量映射到自己的内部数据结构,再执行相应的操作。

MIB 的设计比较灵活,对象和网络管理通信协议相对独立,只要有需要就可以定义新的 MIB 变量并标准化,而不需要改变协议。各厂商设计了新的网络设备或新的网络协议时,可以自行定义相应的 MIB 变量,对这些新的网络设备或新的网络协议进行管理。事实上,目前已定义了许多新的 MIB 变量。例如,以太网接口的 MIB、网桥的 MIB、FDDI 的 MIB、PPP 的 IP 网络控制协议 MIB 等。

MIB 只是定义了管理对象的组织结构,使其他系统可以知道如何访问各个管理对象。对于代理如何收集 MIB 中管理对象的数据以及怎样使用这些数据,MIB 不做规定。

2. 管理信息结构

管理信息结构(Structure of Management Information,SMI)是一套规则,规定如何命名管理对象、如何定义管理对象。所有的管理对象分层次地按树形结构进行组织。某个对象的名称反映它在这个树形结构中的位置,标明了在 MIB 中通过怎样的路径可以访问到这个对象。如图 10.2 所示是这棵树顶部的一部分。

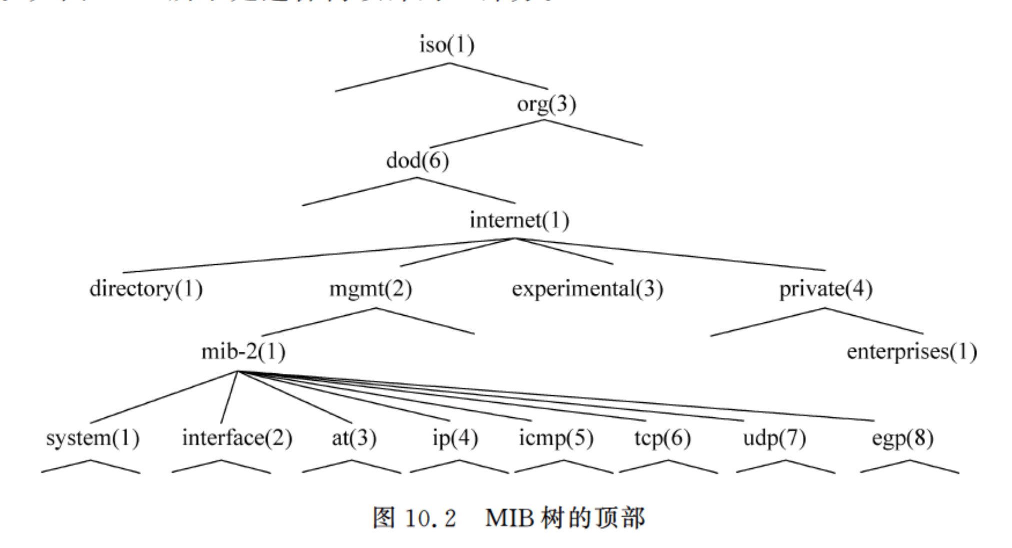


图 10.2　MIB 树的顶部

SMI 采用 OSI 的抽象语法表示(Abstract Syntax Notation One,ASN.1)的 Object Identifier 类型对 MIB 的管理对象进行命名。例如,图 10.2 中的 ip 对象命名为{iso(1)org(3)dod(6)internet(1)mgmt(2)mib-2(1)ip(4)}或 1.3.6.1.2.1.4。

3. 简单网络管理协议

简单网络管理协议的主要设计思想是协议应尽可能简单,基本功能是监视网络性能、检测分析网络差错和配置网络设备。由于 SNMP 的设计是基于互联网协议的用户数据报协议 UDP 之上的,所以 SNMP 提供的是一种面向无连接服务,它不能确保其他实体一定能收到管理信息流。

10.2.2　SNMP

SNMP 是 TCP/IP 网络的网络管理协议,现在已被广大厂商接受,成为一种事实上的标准。随着 SNMP 的广泛使用,已经从 SNMPv1(第 1 版)发展到 SNMPv2(第 2 版)和 SNMPv3(第 3 版)。

SNMP 规定了 5 种协议数据单元(PDU),即 SNMP 报文,用来在管理进程和代理之间进行交换,其中包括以下几项。

(1) get-request 操作:从代理进程处提取一个或多个进程值。

(2) get-next-request 操作:从代理进程处提取紧跟当前参数值的下一个参数值。

(3) set-request 操作:设置代理进程一个或多个参数值。

(4) get-response 操作:返回一个或多个参数值。此操作是由代理进程发出的,是前面 3 种操作的响应操作。

(5) trap 操作:代理进程主动发出的报文,通知管理进程有某些事情发生。

前 3 种操作是由管理进程向代理进程发出的，后两个操作是代理进程向管理进程发出的。SNMP 这 5 种报文的操作如图 10.3 所示。在代理进程端使用 161 端口来接收 get 或 set 报文，在管理进程端使用 162 端口来结束 trap 报文。

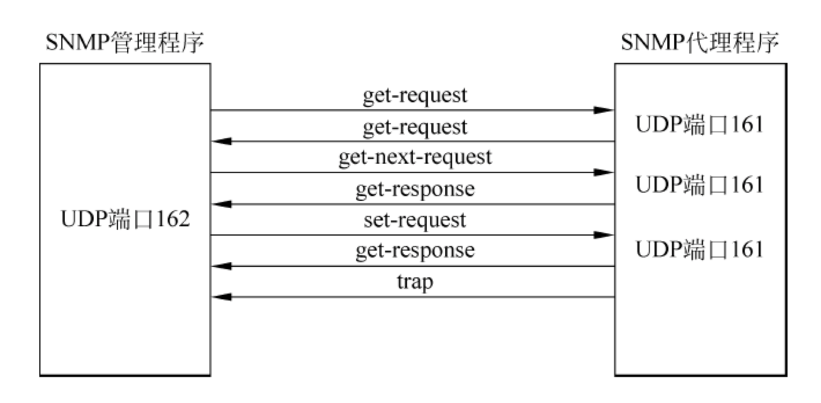


图 10.3　SNMP 的 5 种报文操作

封装成 UDP 数据报的 SNMP 报文格式如图 10.4 所示。从中可以看出，一个 SNMP 报文共分成 3 个部分：公共 SNMP 头部、get/set 头部、trap 头部和变量绑定。

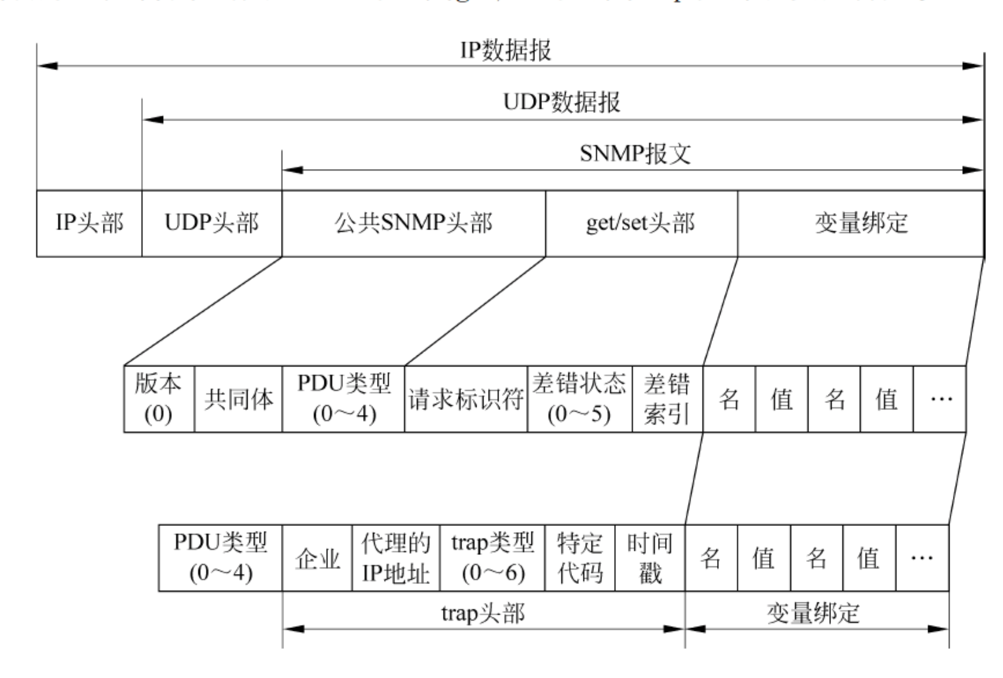


图 10.4　SNMP 报文格式

(1) 公共 SNMP 头部共有 3 个字段。

① 版本：写入本字段的值是版本号减 1。对于 SNMPv1，则应写入 0。

② 共同体：是一个字符串，作为管理进程和代理进程之间的明文口令，通常使用的 6 个字符是“p”“u”“b”“l”“i”“c”。

③ PDU 类型：可填入 0～4 中的一个数字，其对应名称如表 10.2 所示。

(2) get/set 头部共有 3 个字段。

① 请求标识符(Request ID)：是由管理进程设置的一个整数值。代理进程在发送 get-response 报文时也要返回此请求标识符。管理进程可同时向许多代理发出 get 报文，这些

报文都使用 UDP 传送，先发送的有可能后到达。请求标识符可使管理进程识别返回的响应报文对应于哪一个请求报文。

表 10.2 PDU 类型

PDU 类型	名 称
0	get-request
1	get-next-request
2	get-response
3	set-request
4	trap

② 差错状态(Error Status)：由代理进程应答时填写 0～5 中的一个数，如表 10.3 所示。

表 10.3 差错状态描述

差错状态	名 字	说 明
0	noError	一切正常
1	tooBig	代理无法将回答装入一个 SNMP 报文之中
2	noSuchName	操作指明了一个不存在的变量
3	badValue	一个 set 操作指明了一个无效值或无效语法
4	readOnly	管理进程试图修改一个只读变量

③ 差错索引(Error Index)：当出现 noSuchName、badValue 或 readOnly 差错时，由代理进程在应答时设置一个整数，指明有差错的变量在变量列表中的偏移。

(3) trap 头部有以下几项。

① 企业(Enterprise)：填入 trap 报文的网络设备的对象标识符。此对象标识符在如图 10.2 所示的对象命名树的 enterprises 节点下面的一棵子树上。

② trap 类型：此字段的名称是 generic-trap，共分为 6 种类型，如表 10.4 所示。

当使用上述类型 2、3 和 5 时，在报文后面变量部分的第一个变量应标识响应的接口。

③ 特定代码(Specific-Code)：如果 trap 类型为 6，指明代理自定义的时间，否则为 0。

④ 变量绑定(Variable-Bindings)：指明一个或多个变量的名和对应的值。在 get 和 get-next 报文中，变量的值应忽略。

表 10.4 trap 的 6 种类型

差错状态	名 字	说 明
0	coldStart	代理进行了初始化
1	warmStart	代理进行了重新初始化
2	linkDown	一个接口从工作状态变为故障状态
3	linkUp	一个接口从故障状态变为工作状态
4	authenticationFailure	从 SNMP 管理进程接收到具有一个无效共同体的报文
5	egpNeighborLoss	一个 EGP 相邻路由器变为故障状态

10.3 网络安全

10.3.1 网络安全概述

1. 网络安全的定义

以 Internet 为代表的全球性信息化浪潮日益高涨，信息网络技术的应用正日益普及和广泛，应用层次正在深入，应用领域从传统的、小型业务系统逐渐向大型、关键业务系统扩展，典型的如党政部门信息系统、金融业务系统、企业商务系统等。伴随着网络的普及，安全日益成为影响网络效能的重要问题，而 Internet 所具有的开放性、国际性和自由性在增加应用自由度的同时，对安全提出了更高的要求，网络的安全属性主要表现在以下几个方面。

(1) 保密性(Secrecy)：信息不泄露给非授权的用户、实体或进程。

(2) 完整性(Integrity)：信息在存储或传输过程中保持不被修改、不被破坏和丢失的特性。

(3) 可用性(Available)：可被授权实体访问并按需求使用的特性。

(4) 真实性(Authenticity)(认证性、不可抵赖性)：在信息交互过程中，确信参与者的真实同一性，所有参与者都不能否认和抵赖曾经完成的操作和承诺。

(5) 可控性(Controllable)：对信息的传播路径、范围及其内容所具有的控制能力。

从狭义的保护角度来看，计算机网络安全是指计算机及其网络系统资源和信息资源不受自然和人为有害因素的威胁与危害，即指计算机与网络系统的硬件、软件及其系统中的数据受到保护，不因偶然的或者恶意的原因而遭到破坏、更改、泄露，保障系统能连续可靠地运行。计算机安全从本质上来讲是系统的信息安全。

从广义的保护角度来看，凡是涉及计算机网络上信息的保密性、完整性、可用性、真实性和可控性的相关技术与理论都是计算机网络安全的研究领域。它涵盖了与计算机网络系统有关的所有硬件、软件、数据、管理、环境内容等内容。通常所说的网络安全主要是指狭义的网络安全。

2. 网络安全的衡量

为了衡量计算机网络(系统)的安全，世界各国制定了很多的网络安全标准，其中最常见的是美国可信计算机安全评价标准(TCSEC)，它由美国国防部开发。TCSEC 将网络(系统)安全分为 4 级，具体介绍如下。

1) D 级

D 级是最低的安全级别，对系统提供最小的安全防护。系统的访问控制没有限制，无须登录就可以访问数据，这个级别典型的例子有 MS-DOS 和 Windows 98 等。

2) C 级

属于自由选择性安全保护，在设计上有自我保护和审计功能，可对主体行为进行审计与约束，其安全策略主要是自主存取控制，可实现数据保护，确保非授权用户无法访问、对存取权限的传播进行控制及个人用户数据的安全管理。C 级又分成 C1 和 C2 两个安全子级。

(1) C1 级称为选择性保护级(Discretionary Security Protection)。它可以实现自主安

全保护,对用户和数据的分离,保护或限制用户权限的传播。

(2) C2：具有访问控制环境的权力,比 C1 的访问控制划分得更为详细,能够实现受控安全保护、个人账户管理、审计和资源隔离。这个级别的系统包括 UNIX、Linux 和 Windows NT 操作系统。

3) B 级

能够提供强制性安全保护和多级安全,可以实现自主存取控制和强制存取控制,通常包括所有对象都有标识、安全标识对普通用户是不可变更的以及相应的审计功能。B 级包括 B1、B2、B3 三个安全子级。

(1) B1：称为标识安全保护(Labeled Security Protection),它是支持多级安全的第一个级别。这一级别可以使一个处于强制性访问控制之下的对象,不允许文件的拥有者改变其权限。

(2) B2：称为结构保护级别(Structured Protection),要求所有对象都有安全标签以实现低级别的用户不能访问敏感信息。

(3) B3：称为安全域保护级别(Security Domain Protection),使用安装硬件的方式来加强域的安全。

4) A 级

A 级又被称为验证设计级(Verify Design),是目前最高的安全级别。在 A 级中,安全的设计必须给出形式化设计说明和验证,需要有严格的数学推导过程,同时应该包含对加密信道和可信分布的分析,即需要确保系统的部件来源有安全保证,以避免出现安全隐患。

10.3.2　网络安全风险

网络安全风险主要有 4 种基本的安全威胁：信息泄露、完整性破坏、拒绝服务、非法使用。主要的可实现的威胁包括：渗入威胁,如假冒、旁路、授权侵犯；植入威胁,如特洛伊木马、陷门。目前,计算机互联网面临的安全威胁表现形式主要有以下几个方面。

1. 非法访问和破坏

非授权访问是指没有预先经过同意就使用网络或计算机资源,如有意避开系统访问控制机制,对网络设备及资源进行非正常使用,或擅自扩大权限,越权访问信息。主要有以下几种形式：假冒、身份攻击、非法用户进入网络系统进行违法操作、合法用户以未授权方式进行操作等。操作系统总不免存在这样那样的漏洞,一些人就利用系统的漏洞进行网络攻击,其主要目标就是对系统数据的非法访问和破坏(如“黑客”攻击)。“黑客”攻击已有十几年的历史,黑客活动几乎覆盖了所有的操作系统,包括 UNIX、Windows NT、VM、VMS 以及 MVS。

2. 拒绝服务攻击

拒绝服务攻击是一种破坏性攻击,最早的拒绝服务攻击(Denial Of Service Attack)是“电子邮件炸弹”,它能使用户在很短的时间内收到大量电子邮件,使用户系统不能处理正常工作,严重时会使系统崩溃、网络瘫痪。它不断对网络服务系统进行干扰,改变其正常的作业流程,执行无关程序使系统响应减慢甚至瘫痪,影响正常用户的使用,甚至使合法用户被排斥而不能进入计算机网络系统或不能得到相应的服务。

3. 计算机病毒

计算机病毒程序很容易做出，有着巨大的破坏性，其危害已被人们所认识。单机病毒就已经让人们“谈毒色变”了，而通过网络传播的病毒，无论是在传播速度、破坏性，还是在传播范围等方面都是单机病毒不能比拟的。

4. 特洛伊木马

特洛伊木马(Trojan Horse)的名称来源于古希腊的历史故事。特洛伊程序一般是由编程人员编制，它提供了用户所不希望的功能，这些额外的功能往往是有害的。把预谋的、有害的功能隐藏在公开的功能中，以掩盖其真实企图。

5. 破坏数据完整性

破坏数据完整性是指以非法手段窃得对数据的使用权，删除、修改、插入或重发某些重要信息，修改网络上传输的数据，以及销毁网络上传输的数据，替代网络上传输的数据，重复播放某个分组序列，改变网络上传输的数据包的先后次序，使攻击者获益，以干扰用户的正常使用。

6. 蠕虫

蠕虫(Worms)是一个或一组程序，它可以从一台机器向另一台机器传播。它同病毒不一样，它不需要修改宿主程序就能传播。

7. 陷门

陷门(Trap Doors)是指为攻击者提供“后门”的一段非法的操作系统程序。这一般是指一些内部程序人员为了特殊的目的，在所编制的程序中潜伏代码或保留漏洞。

8. 隐蔽通道

一种允许以违背合法的安全策略的方式进行操作系统进程间通信(IPC)的通道，它分为隐蔽存储通道和隐蔽时间通道。隐蔽通道的重要参数是带宽。

9. 信息泄露或丢失

信息泄露或丢失是指敏感数据在有意或无意中被泄露出去或丢失，通常包括：信息在传输中丢失或泄露(如“黑客”们利用电磁泄露或搭线窃听等方式截获机密信息，或通过对信息流向、流量、通信频度和长度等参数的分析，推出有用信息，如用户口令、账号等)，信息在存储介质中丢失或泄露，通过建立隐蔽隧道等窃取敏感信息等。

在所有的操作系统中，由于 UNIX 操作系统的核心代码是公开的，这使其成为最易受攻击的目标。攻击者可能先设法登录到一台 UNIX 的主机上，通过操作系统的漏洞来取得特权，然后再以此为据点访问其余主机，这被称为“跳跃”(I Stand-hopping)。攻击者在到达目的主机之前往往会先经过几次这种跳跃。这样，即使被攻击网络发现了攻击者从何处发起攻击，管理员也很难顺次找到他们的最初据点，而且他们在窃取某台主机的系统特权后，在退出时会删掉系统日志。用户只要能登录到 UNIX 操作系统上，就能相对容易地成为超级用户。所以，如何检测系统自身的漏洞，保障网络的安全，已成为一个日益紧迫的问题。

10.3.3 网络安全策略

安全策略是指在一个特定的环境里，为保证提供一定级别的安全保护所必须遵守的规则。网络安全策略包括对企业的各种网络服务的安全层次和用户的权限进行分类，确定管

理员的安全职责，如何实施安全故障处理、网络拓扑结构、入侵及攻击的防御和检测、备份和灾难恢复等内容。在本书中所说的安全策略主要是指系统安全策略，主要涉及4个方面：物理安全策略、访问控制策略、信息加密策略和网络安全管理策略。

1. 物理安全策略

物理安全策略的目的是保护计算机系统、网络服务器、打印机等硬件实体和通信链路免受自然灾害、人为破坏和搭线攻击；验证用户的身份和使用权限，防止用户越权操作；确保计算机系统有一个良好的电磁兼容工作环境；建立完备的安全管理制度，防止非法进入计算机控制室和各种偷窃、破坏活动的发生。

抑制和防止电磁泄露（TEMPEST技术）是物理安全策略的一个主要问题。目前，主要的防护措施有两类。一类是对传导发射的防护，主要是对电源线和信号线加装性能良好的滤波器，减小传输阻抗和导线间的交叉耦合。另一类是对辐射的防护，这类防护措施又可分为以下两种：一是采用各种电磁屏蔽措施，如对设备的金属屏蔽和各种接插件的屏蔽，同时对机房的下水管、暖气管和金属门窗进行屏蔽与隔离；二是干扰的防护措施，即在计算机系统工作的同时，利用干扰装置产生一种与计算机系统辐射相关的伪噪声向空间辐射来掩盖计算机系统的工作频率和信息特征。

2. 访问控制策略

访问控制是网络安全防范和保护的主要策略，它的主要任务是保证网络资源不被非法使用和非常访问。它也是维护网络系统安全、保护网络资源的重要手段。各种安全策略必须相互配合才能真正起到保护作用，但访问控制可以说是保证网络安全非常重要的核心策略之一。

1）入网访问控制

入网访问控制为网络访问提供了第一层访问控制。它控制哪些用户能够登录到服务器并获取网络资源，控制准许用户入网的时间和准许他们在哪台工作站入网。

用户的入网访问控制可分为3个步骤：用户名的识别与验证、用户口令的识别与验证、用户账号的默认限制检查。三道关卡中只要任何一关未过，该用户便不能进入该网络。

对网络用户的用户名和口令进行验证是防止非法访问的第一道防线。用户注册时首先输入用户名和口令，服务器将验证所输入的用户名是否合法。如果验证合法，才能继续验证用户输入的口令；否则，用户将被拒之网络之外。用户口令是用户入网的关键所在。为保证口令的安全性，用户口令不能显示在显示屏上，口令长度应不少于6个字符，口令字符最好是数字、字母和其他字符的混合，用户口令必须经过加密，加密的方法很多，其中最常见的方法包括：基于单向函数的口令加密；基于测试模式的口令加密；基于公钥加密方案的口令加密；基于平方剩余的口令加密；基于多项式共享的口令加密；基于数字签名方案的口令加密等。经过上述方法加密的口令，即使是系统管理员也难以得到它。用户还可采用一次性用户口令，也可用便携式验证器（如智能卡）来验证用户的身份。

网络管理员应该可以控制和限制普通用户的账号使用、访问网络的时间和方式。用户名或用户账号是所有计算机系统中最基本的安全形式。用户账号应只有系统管理员才能建立。用户口令应是每个用户访问网络所必须提交的“证件”，用户可以修改自己的口令，但系统管理员应该可以控制口令的几个方面的限制：最小口令长度、强制修改口令的时间间隔、口令的唯一性、口令过期失效后允许入网的宽限次数。

用户名和口令验证有效之后，再进一步履行用户账号的默认限制检查。网络应能控制用户登录入网的站点、限制用户入网的时间、限制用户入网的工作站数量。当用户对交费网络的访问“资费”用尽时，网络还应能对用户的账号加以限制，用户此时应无法进入网络访问网络资源。网络应对所有用户的访问进行审计。如果多次输入口令不正确，则认为是非法用户的入侵，应给出报警信息。

2）网络权限控制

网络权限控制是针对网络非法操作所提出的一种安全保护措施。用户和用户组被赋予一定的权限。网络控制用户和用户组可以访问哪些目录、子目录、文件和其他资源。可以指定用户对这些文件、目录、设备能够执行哪些操作。受托者指派和继承权限屏蔽(IRM)可作为其两种实现方式。受托者指派控制用户和用户组如何使用网络服务器的目录、文件和设备。继承权限屏蔽相当于一个过滤器，可以限制子目录从父目录那里继承哪些权限。可以根据访问权限将用户分为以下几类。

(1) 特殊用户(即系统管理员)。

(2) 一般用户，系统管理员根据他们的实际需要为他们分配操作权限。

(3) 审计用户，负责网络的安全控制与资源使用情况的审计。用户对网络资源的访问权限可以用一张访问控制表来描述。

3）目录级安全控制

网络应允许控制用户对目录、文件、设备的访问。用户在目录一级指定的权限对所有文件和子目录有效，用户还可进一步指定对目录下的子目录和文件的访问权限。对目录和文件的访问权限一般有 8 种：系统管理员权限(Supervisor)、读权限(Read)、写权限(Write)、创建权限(Create)、删除权限(Erase)、修改权限(Modify)、文件查找权限(File Scan)、存取控制权限(Access Control)。用户对文件或目标的有效权限取决于以下 3 个因素：用户的委托者指派、用户所在组的委托者指派、继承权限屏蔽取消的用户权限。一个网络系统管理员应当为用户指定适当的访问权限，这些访问权限控制着用户对服务器的访问。8 种访问权限的有效组合可以让用户有效地完成工作，同时又能有效地控制用户对服务器资源的访问，从而加强了网络和服务器的安全性。

4）属性安全控制

当使用文件、目录和网络设备时，网络系统管理员应给文件、目录等指定访问属性。属性安全控制可以将给定的属性与网络服务器的文件、目录和网络设备联系起来。属性安全在权限安全的基础上提供更进一步的安全性。网络上的资源都应预先标出一组安全属性。用户对网络资源的访问权限对应一张访问控制表，用以表明用户对网络资源的访问能力。属性设置可以覆盖已经指定的任何委托者指派和有效权限。属性往往能控制以下几个方面的权限：向某个文件写数据、复制一个文件、删除目录或文件、查看目录和文件、执行文件、隐含文件、共享、系统属性等。网络的属性可以保护重要的目录和文件，防止用户对目录和文件的误删除、修改、显示等。

5）网络服务器安全控制

网络允许在服务器控制台上执行一系列操作。用户使用控制台可以装载和卸载模块，可以安装和删除软件等。网络服务器安全控制包括：设置口令锁定服务器控制台，以防止非法用户修改、删除重要信息或破坏数据；设定服务器登录时间限制、非法访问者检测和关

闭的时间间隔。

6）网络监测和锁定控制

网络管理员应对网络实施监控，服务器应记录用户对网络资源的访问，对非法的网络访问，服务器应以图形、文字或声音等形式报警，以引起网络管理员的注意。如果不法之徒试图进入网络，网络服务器应会自动记录企图尝试进入网络的次数，如果非法访问的次数达到设定数值，那么该账户将被自动锁定。

7）网络端口和节点的安全控制

网络中的服务器端口往往使用自动回复设备、静默调制解调器加以保护，并以加密的形式来识别节点的身份。自动回复设备用于防止假冒合法用户，静默调制解调器用以防范黑客的自动拨号程序对计算机进行攻击。网络还常对服务器端和用户端采取控制，用户必须携带证实身份的验证器（如智能卡、磁卡、安全密码发生器）。在对用户的身份进行验证之后，才允许用户进入用户端。然后，用户端和服务器端再进行相互验证。

8）防火墙控制

防火墙是近期发展起来的一种保护计算机网络安全的技术性措施，它是一个用以阻止网络中的黑客访问某个机构网络的屏障，也可称为控制进/出两个方向通信的门槛。在网络边界上通过建立起来的相应网络通信监控系统来隔离内部网和外部网，以阻止外部网的侵入。

3. 信息加密策略

信息加密的目的是保护网内的数据、文件、口令和控制信息，保护网上传输的数据。网络加密常用的方法有链路加密、端点加密和节点加密 3 种。链路加密的目的是保护网络节点之间的链路信息安全；端点加密的目的是对源端用户到目的端用户的数据提供保护；节点加密的目的是对源节点到目的节点之间的传输链路提供保护。用户可根据网络情况酌情选择上述加密方式。

信息加密过程是通过形形色色的加密算法来具体实现的，它以很小的代价提供很大的安全保护。在多数情况下，信息加密是保证信息机密性的唯一方法。据不完全统计，到目前为止，已经公开发表的各种加密算法多达数百种。如果按照收发双方密钥是否相同来分类，可以将这些加密算法分为常规密码算法和公钥密码算法。

在常规密码中，收信方和发信方使用相同的密钥，即加密密钥和解密密钥是相同或等价的。比较著名的常规密码算法有 DES、Triple DES、GDES、New DES、IDEA、RC4、RC5 以及以代换密码和转轮密码为代表的古典密码等。在众多的常规密码中影响最大的是 DES 密码。

常规密码的优点是有很强的保密强度，且经受住时间的检验和攻击，但其密钥必须通过安全的途径传送。因此，其密钥管理成为系统安全的重要因素。

在公钥密码中，收信方和发信方使用的密钥互不相同，而且几乎不可能从加密密钥推导出解密密钥。比较著名的公钥密码算法有 RSA、Diffe-Hellman、Rabin、Ong-Fiat-Shamir、零知识证明的算法、椭圆曲线、EIGamal 算法等。最有影响的公钥密码算法是 RSA，它能抵抗目前为止已知的所有密码攻击。

公钥密码的优点是可以适应网络的开放性要求，且密钥管理问题也较为简单，尤其可方便地实现数字签名和验证。但其算法复杂，加密数据的速率较低。尽管如此，随着现代电子技术和密码技术的发展，公钥密码算法将是一种很有前途的网络安全加密体制。

当然在实际应用中人们通常将常规密码和公钥密码结合在一起使用，例如，利用 DES 或者 IDEA 来加密信息，而采用 RSA 来传递会话密钥。如果按照每次加密所处理的比特来分类，可以将加密算法分为序列密码和分组密码。前者每次只加密一个比特而后者则先将信息序列分组，每次处理一个组。

密码技术是网络安全非常有效的技术之一。一个加密网络，不但可以防止非授权用户的搭线窃听和入网，而且也是对付恶意软件的有效方法之一。

4. 网络安全管理策略

在网络安全中，除了采用上述技术措施之外，加强网络的安全管理，制定有关规章制度，对于确保网络安全、可靠地运行，也将起到十分有效的作用。安全管理策略是指在一个特定的环境里，为提供一定级别的安全保护所必须遵守的规则。该策略模型包括建立安全环境的如下 3 个重要组成部分。

（1）威严的法律：安全的基石是社会法律、法规与手段，这是建立一套安全管理的标准和方法。即通过建立与信息安全相关的法律、法规，使非法分子慑于法律，不敢轻举妄动。

（2）先进的技术：先进的安全技术是信息安全的根本保障，用户对自身面临的威胁进行风险评估，根据安全服务的种类，选择相应的安全机制，然后集成先进的安全技术。

（3）严格的管理：制定有关网络操作使用规程和人员出入机房的管理制度；制定网络系统的维护制度和应急措施等。各网络使用机构、企业和单位应建立相宜的信息安全管理办法，加强内部管理，建立审计和跟踪体系，提高整体的信息安全意识。

网络安全管理策略包括：确定安全管理等级和安全管理范围；制定有关网络操作使用规程和人员出入机房的管理制度；制定网络系统的维护制度和应急措施等。

10.4 数据加密和数字证书

在信息时代，许多数据是十分重要的。信息可以帮助团体或个人，使他们受益，同样，信息也可以用来对他们构成威胁，造成破坏。在竞争激烈的大公司中，商业间谍经常会获取对方的情报。因此，就需要一种强有力的安全措施来保护机密数据不被窃取或篡改。数据加密与解密从宏观上讲是非常简单的，也很容易理解。加密与解密的一些方法是非常直接的，很容易掌握，可以很方便地对机密数据进行加密和解密。

10.4.1 数据加密技术

作为保障数据安全的一种方式，数据加密起源于公元前 2000 年。埃及人是最先使用特别的象形文字作为信息编码的人。随着时间的推移，巴比伦、美索不达米亚和希腊文明都开始使用一些方法来保护他们的书面信息。加密技术被 Julius Caesar（恺撒大帝）使用，也曾用于历次战争中，包括美国独立战争、美国内战和两次世界大战。最广为人知的编码机器是 German Enigma，在第二次世界大战中德国人利用它创建了加密信息。此后，由于 Alan Turing 和 Ultra 计划以及其他人的努力，终于对德国人的密码进行了破解。当初，计算机的研究就是为了破解德国人的密码，当时人们并没有想到计算机会给今天带来的信息革命。

随着计算机的发展、运算能力的增强，过去的密码都变得十分简单了，于是人们又不断地研究出了新的数据加密方式，如私有密钥算法和公共密钥算法。可以说，是计算机推动了数据加密技术的发展。

因特网一方面是危险的，而且这种危险是 TCP/IP 所固有的，一些基于 TCP/IP 的服务也是极不安全的；另一方面，因特网给众多的商家带来了无限的商机，因为因特网把全世界连在了一起，走向因特网就意味着走向了世界。为了使因特网变得安全和充分利用其商业价值，人们选择了数据加密和基于加密技术的身份认证。

加密在网络上的作用就是防止有价值的信息在网络上被拦截和窃取。一个简单的例子就是密码的传输。计算机密码极为重要，许多安全防护体系是基于密码的，密码的泄露会导致安全体系的全面崩溃。通过网络进行登录时，所输入的密码以明文的形式被传输到服务器，而在网络上窃听是一件极为容易的事情，所以很有可能黑客会嗅探并窃得用户的密码，如果是 Root 用户或 Administrator 用户，那么后果将是极为严重的：网络上的数据被嗅探和劫持。

解决的方法就是加密，加密后的口令即使被黑客获得也是不可读的，除非加密密钥或加密方式十分脆弱，黑客破解。不管怎样，加密的使用使黑客不会轻易获得口令。

身份认证是基于加密技术的，它的作用就是确定用户是否是真实的。简单的例子就是电子邮件，当用户收到一封电子邮件时，邮件上面标有发信人的姓名和信箱地址，很多人可能会简单地认为发信人就是信上说明的那个人，但实际上伪造一封电子邮件是件极为容易的事。在这种情况下，用户需要用电子邮件源身份认证技术来防止电子邮件伪造，这样就有理由使用户确信写信的人就是信头上说明的人，有些站点提供入站 FTP 和 WWW 服务，当然用户通常接触的这类服务是匿名服务，用户的权利要受到限制，但也有的服务不是匿名的，如公司为了信息交流为用户的合作伙伴提供非匿名的 FTP 服务，或开发小组把他们的 Web 网页上传到用户的 WWW 服务器上，现在的问题就是，用户如何确定正在访问用户的服务器的人就是用户认为的那个人，身份认证是一个好的解决方案。

有些时候，用户可能需要对一些机密文件进行加密，不一定因为要在网络间传输该文件，而是担心有人窃得计算机密码而获得该机密文件，对文件实行加密，从而实现多重保护显然会使用户感到安心，例如，在 UNIX 操作系统中可以用 crypt 命令对文件进行加密，尽管这种加密手段已不是那么先进，甚至有被破解的较大可能性。

1. 对称密码技术

在传统的加密算法中，加密密钥与解密密钥是相同的，或者可以由其中一个推知另一个，称为对称密钥加密算法。这样的密钥必须秘密保管，只能为授权用户所知，授权用户既可以用该密钥加密信息，也可以用该密钥解密信息。

最著名的保密密钥或对称密钥加密算法 DES(Data Encryption Standard)是由 IBM 公司在 20 世纪 70 年代发展起来的，并经美国政府的加密标准筛选后，于 1976 年 11 月被美国政府采用，DES 随后被美国国家标准局和美国国家标准协会(American National Standard Institute，ANSI)承认。

DES 使用 56b 密钥对 64b 的数据块进行加密，并对 64b 的数据块进行 16 轮编码。进行每轮编码时，一个 48b 的“每轮”密钥值由 56b 的完整密钥得出来。DES 用软件进行解码需用很长时间，而用硬件解码的速度非常快。幸运的是，当时大多数黑客并没有足够的设备

制造出这种硬件设备。在1977年,人们估计要耗资两千万美元才能建成一台专门计算机用于DES的解密,而且需要12h的破解才能得到结果。当时DES被认为是一种十分“强壮”的加密方法。

但是,当今的计算机速度越来越快了,制造一台这样特殊机器的花费已经降到了10万美元左右,而用它来保护10亿美元的银行间线缆时,就要从新考虑了。另外,如果只用它来保护一台服务器,那么DES确实是一种好的办法,因为黑客绝不会仅仅为入侵一台服务器而花那么多的钱破解DES密文。因为现在已经能用20万美元制造一台破译DES的特殊计算机,所以现在对要求“强壮”加密的场合已经不再适用了。

确定一种新的加密法是否真的安全是极为困难的,何况DES的密码学缺点只是密钥长度相对比较短,所以人们并没有放弃使用DES,而是想出了一个解决其长度的方法,即采用三重DES。这种方法是用两个密钥对明文进行3次加密,假设两个密钥是K1和K2。

(1) 用密钥K1进行DES解密。

(2) 用K2对步骤(1)的结果进行DES解密。

(3) 使用密钥K1对步骤(2)的结果进行DES加密。

这种方法的缺点是,花费是原来的3倍,但从另一方面来看,三重DES的112b密钥长度是很“强壮”的加密方式了。

DES的保密性仅取决于对密钥的保密,而算法是公开的。DES内部的复杂结构是至今没有找到捷径破译方法的根本原因。

2. 非对称密码技术

非对称密钥密码算法又叫公开密钥密码算法,公开密钥密码体制最主要的特点就是加密和解密使用不同的密钥,每个用户保存着一对密钥——公开密钥PK和秘密密钥SK,因此,这种体制又称为双钥或非对称密钥密码体制。在公钥加密算法下,公钥是公开的,任何人可以用公钥加密信息,再将密文发送给私钥拥有者;私钥是保密的,用于解密其接收的公钥加密过的信息。典型的公钥加密算法如RSA,是目前使用比较广泛的加密算法。在互联网上通过浏览器进行的数据安全传输,如Netscape Navigator和Microsoft Internet Explorer都使用了该算法。RSA加密算法建立在大数因子分解的复杂性上。RSA的保密性在于大数的分解难度上,如果大数分解成功,则RSA也就无保密性可言了。一个好的加密算法的重要特点之一是具有这种能力:可以定一个密码或密钥,并用它来加密明文,不同的密码或密钥产生不同的密文。这又分为两种方式:对称密钥密码算法和非对称密钥密码算法。所谓对称密钥密码算法就是加密解密都使用相同的密钥,非对称密钥密码算法就是加密解密使用不同的密钥。非常著名的PGP公钥加密以及RSA加密方法都是非对称密钥密码算法。加密密钥,即公钥,与解密密钥,即私钥,是非常不同的。从数学理论上讲,几乎没有真正不可逆的算法存在。例如,对于一个输入a执行一个操作得到结果b,那么可以基于b,做一个相对应的操作,导出输入a。在一些情况下,对于每一种操作,可以得到一个确定的值,或者该操作没有定义(例如,除数为0)。对于一个没有定义的操作来讲,基于加密算法,可以成功地防止把一个公钥变换成为私钥。因此,要想破译非对称密钥密码算法,找到那个唯一的密钥,唯一的方法只能是反复地试验,而这需要大量的处理时间。

RSA加密算法使用了两个非常大的素数来产生公钥和私钥。即使通过因数分解从一个公钥可以得到私钥,但这个运算所包含的计算量是非常巨大的,以至于实际上是不可行

的。加密算法本身也是很慢的，这使得使用 RSA 加密算法加密大量的数据变得有些不可行。这就使得一些现实中的加密算法都基于 RSA 加密算法。PGP 算法（以及大多数基于 RSA 算法的加密方法）使用公钥来加密一个对称加密算法的密钥，然后再利用一个快速的对称加密算法来加密数据。这个对称算法的密钥是随机产生的，是保密的，因此，得到这个密钥的唯一方法就是使用私钥来解密。

举一个例子：假定现在要使用密钥 12345 加密一些数据。利用 RSA 公钥，使用 RSA 算法加密这个密钥 12345，并把它放在要加密的数据的前面（可能后面跟着一个分隔符或文件长度，以区分数据和密钥），然后，使用对称加密算法加密正文，使用的密钥就是 12345。当对方收到时，解密程序找到加密过的密钥，并利用 RSA 私钥解密出来，然后再确定出数据的开始位置，利用密钥 12345 来解密数据。这样就使得一个可靠的经过高效加密的数据可以安全地传输和解密。

10.4.2　数字证书和公钥基础设施

随着 Internet 上各类应用的发展，尤其是电子商务应用的发展，为保证商务、交易及支付活动的真实可靠，需要有一种机制来验证活动中各方的真实身份。安全认证是维持电子商务活动正常进行的保证。PKI 公开密钥体系就是一种基于加密技术的安全认证机制。

PKI（Public Key Infrastructure，公钥基础设施）能够解决安全隐私系统和数据的问题，可以用于网络安全和保护数据。它基于一种非对称密钥的密码理论。用户利用 PKI 平台提供的安全服务进行安全通信，网上进行的任何需要安全服务的通信都建立在公钥的基础之上，而与公钥相对的私钥只掌握在它们与之通信的另一方，通过数字签名和电子（数字）证书的使用，确保传输的电子文件的完整性、真实性和不可抵赖性。

为了理解 PKI，下面从以下几个方面进行描述。

1. 数据完整性验证

消息的发送者用要发送的消息和一定的算法生成一个附件，并将附件与消息一起发送出去；消息的接收者收到消息和附件后，用同样的算法与接收到的消息生成一个新的附件；将新的附件与接收到的附件进行比较，如果相同，则说明收到的消息是正确的；否则说明消息在传送中出现了错误。

在算法中用到的数学函数叫作单向散列函数（One-Way Hash Function），也叫压缩函数、收缩函数，它是现代密码学的中心，是许多协议的另一个结构模块。单向散列函数长期以来一直在计算机科学中使用，单向散列函数是把可变长度的输入串（叫作预映射，Pre-Image）转换成固定长度的输出串（叫作散列值）的一种函数。

利用单向散列函数生成消息的指纹可以分成两种情况：一种是不带密钥的单向散列函数，在这种情况下，任何人都能验证消息的散列值；另一种是带密钥的单向散列函数，散列值是预映射和密钥的函数，这样只有拥有密钥的人才能验证散列值。单向散列函数的算法实现有很多种，如 Snefru、N-Hash、MD2、MD4、MD5、SHA-1 算法等。

2. 数字签名

数字签名实际上是附加在数据单元上的一些数据或是对数据单元所做的密码变换，这种数据或变换能使数据单元的接收者确认数据单元的来源和数据的完整性，并保护数据，防

止被人(如接收者)伪造。

签名机制的本质特征是该签名只有通过签名者的私有信息才能产生,也就是说,一个签名者的签名只能唯一地由它自己产生。当收发双方发生争议时,第三方(仲裁机构)就能够根据消息上的数字签名来裁定这条消息是否确实由发送方发出,从而实现抗抵赖服务。另外,数字签名应是所发送数据的函数,即签名与消息相关,从而防止数字签名的伪造和重用。

数字签名的实现方法如下。

(1) 使用对称加密和仲裁者实现数字签名。

(2) 使用公开密钥体制进行数字签名。

公开密钥体制的发明,使数字签名变得更简单,它不再需要第三方去签名和验证。签名的实现过程如下。

A 用他的私人密钥加密消息,从而对文件签名。A 将签名的消息发送给 B。B 用 A 的公开密钥解密消息,从而验证签名。

(1) 使用公开密钥体制与单向散列函数进行数字签名。

(2) 加上时间标记的签名:把消息加上时间标记,然后再进行签名,可用在银行支票转账中防止消息与签名一起重用的发生。

(3) 多重签名:对同一消息各自用私有密钥进行多人的签名。

(4) 盲签名:在不可见的消息上进行签名。

3. CA 认证技术

CA(Certificate Authority)即证书机构,是保证公钥的完整性的机构。

(1) 公钥证书(数字证书):在网上进行电子商务活动时,交易双方需要用来表明自己的身份,并使用数字证书来进行交易操作。数字证书中包括证书持有人的身份标识、公钥等信息,并由证书颁发者对证书签字。

(2) 证书的类型与作用如表 10.5 所示。

表 10.5 证书的类型与作用

证书名称	证书类型	主要功能描述
个人证书		用于个人网上交易、网上支付、电子邮件等相关网络作业
单位证书	单位身份证书	用于企事业单位网上交易、网上支付等
	E-mail 证书	用于企事业单位内安全电子邮件通信
	部门证书	用于企事业单位内某个部门的身份认证
服务器证书		用于服务器、安全站点认证等
代码签名证书	个人证书	用于个人软件开发者对其软件的签名
	企业证书	用于软件开发企业对其软件的签名

10.5 防火墙技术

防火墙技术是建立在现代通信网络技术和信息安全技术基础上的应用性安全技术,越来越多地应用于专用网络与公用网络的互联环境之中,尤其以接入 Internet 为最甚。

Internet 的发展给政府机构、企事业单位带来了革命性的改革和开放。通过 Internet，企业可以从异地取回重要数据，但同时又要面对 Internet 开放带来的数据安全的新挑战和新危险，即客户、销售商、移动用户、异地员工和内部员工的安全访问；以及保护企业的机密信息不受黑客和工业间谍的入侵。因此企业必须加筑安全的“战壕”，而这个“战壕”就是防火墙。

10.5.1　防火墙的类型和结构

防火墙是设置在被保护网络和外部网之间的一道屏障，实现网络的安全保护，以防止发生不可预测的、潜在破坏性的侵入。防火墙本身具有较强的抗攻击能力，它是提供信息安全服务、实现网络和信息安全的基础设施。

从防火墙的定义中可以看到防火墙的如下 3 个方面的特征。

(1) 网络位置特性：内部网和外部网之间的所有网络数据流都必须经过防火墙。

(2) 工作原理特性：符合安全策略的数据流才能通过防火墙。

(3) 先决条件：防火墙自身应具有非常强的抗攻击免疫力。

常见防火墙的类型主要有两种：包过滤防火墙和代理防火墙，各有优缺点。

1. 包过滤防火墙

包过滤防火墙工作在 OSI 网络参考模型的网络层和传输层，它根据数据包头的源地址、目的地址、端口号和协议类型等标志确定是否允许通过，如图 10.5 所示。

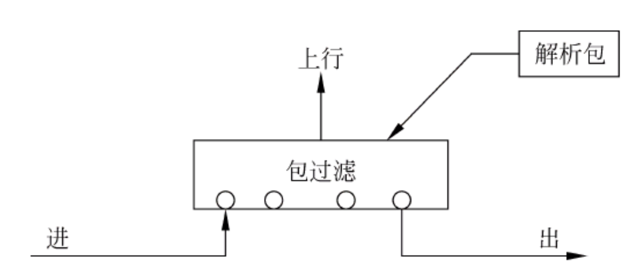


图 10.5　包过滤防火墙的工作方式

数据包过滤是一个网络安全保护机制，它用来控制流出和流入网络的数据。通过控制存在于某一网段的网络流量类型，包过滤可以控制存在于某一网段的服务方式。不符合网络安全的那些服务将被严格限制。基于包中的协议类型和协议字段值，过滤路由器能够区分网络流量；基于协议特定的标准，路由器在其端口能够区分包和限制包的能力叫作包过滤(Packet Filtering)。正是因为这种原因，过滤路由器也可以称为包过滤路由器(Packet Filter Router)。

包过滤的优点如下。

(1) 一个过滤路由器能协助保护整个网络。数据包过滤的主要优点之一是，一台单个的、恰当放置的包过滤路由器有助于保护整个网络。如果仅有一台路由器连接内部网与外部网，不论内部网的大小、内部拓扑结构如何，通过那台路由器进行数据包过滤，在网络安全保护上都会取得较好的效果。

(2) 数据包过滤对用户透明。不像在后面描述的代理(Proxy)，数据包过滤不要求任何自定义软件或者客户机配置，它也不要求用户进行任何特殊的训练或者操作。当数据包过滤路由器决定让数据包通过时，它与普通路由器没什么区别。比较理想的情况是，用户甚至

没有认识到它的存在,除非他们试图做过滤规则中所禁止的事。较强的“透明度”是包过滤的一大优势。

(3) 过滤路由器速度快、效率高。较代理而言,过滤路由器只检查报头相应的字段,一般不查看数据报的内容,而且某些核心部分是由专用硬件实现的,故其转发速率快、效率较高。

包过滤的缺点如下。

(1) 不能彻底防止地址欺骗。大多数包过滤路由器是基于源 IP 地址、目的 IP 地址而进行过滤的。而 IP 地址的伪造是很容易、很普遍的。过滤路由器在这点上大多无能为力。即使按 MAC 地址进行绑定,也是不可信的。对于一些安全性要求较高的网络,过滤路由器是不能胜任的。

(2) 一些应用协议不适合于数据包过滤。即使是完美的数据包过滤实现,也会发现一些协议不很适合于经由数据包过滤安全保护。如 RPC、X-Window 和 FTP。而且,服务代理和 HTTP 的连接,大大削弱了基于源地址和源端口的过滤功能。

(3) 正常的数据包过滤路由器无法执行某些安全策略。数据包过滤路由器上的信息不能完全满足人们对安全策略的需求。例如,数据包说它们来自什么主机(这点还有隐患),而不是什么用户,因此,不能强行限制特殊的用户。同样地,数据包说它到什么端口,而不是到什么应用程序;当通过端口号对高级协议强行进行限制时,不希望在端口上有别的指定协议之外的协议,恶意的知情者能够很容易地破坏这种控制。

数据包工具存在很多局限性。除了各种各样的硬件和软件包普遍具有数据包过滤能力外,数据包过滤仍然算不上是一个完美的工具。许多这样的产品都或多或少地存在局限性,如数据包过滤规则难以配置。

从以上分析可以看出,包过滤防火墙技术虽然能确保一定的安全保护,而且有许多优点,但是包过滤毕竟是第一代防火墙技术,本身存在较多缺陷,不能提供较高的安全性。在实际应用中,现在很少把包过滤技术当作单独的安全解决方案,而是把它与其他防火墙技术糅合在一起使用。

2. 代理防火墙

代理防火墙是一种较新型的防火墙技术,其特点是完全“阻隔”了网络通信流,通过对每种应用服务编制专门的代理程序,实现监视和控制应用层通信流的作用。它分为应用层网关和电路层网关。

应用层网关的工作方式和过滤数据包的防火墙、以路由器为基础的防火墙的工作方式稍有不同。它是基于软件的。

代理防火墙工作于应用层,是针对特定的应用层协议。代理防火墙通过编程来弄清用户应用层的流量,并能在用户层和应用层提供访问控制。而且,还可用来保持一个所有应用程序使用的记录。记录和控制所有进出流量的能力是应用层网关的主要优点之一。那么,代理防火墙是怎样工作的呢?

从图 10.6 中可以看出,代理服务器端作为内部网客户端的服务器拦截住所有要求,也向客户端转发响应。代理客户端(Proxy Client)负责代表内部网客户端向外部网服务器端发出请求,当然也向代理服务器端转发响应。

当某用户(不管是远程的还是本地的)想和一个运行代理的网络建立联系时,此代理(应

用层网关)会阻塞这个连接,然后对连接请求的各个域进行检查。如果此连接请求符合预定的安全策略或规则,代理防火墙便会在用户和服务器之间建立一个“桥”,从而保证其通信。对不符合预定的安全规则的,则阻塞或抛弃。换句话说,“桥”上设置了很多控制。

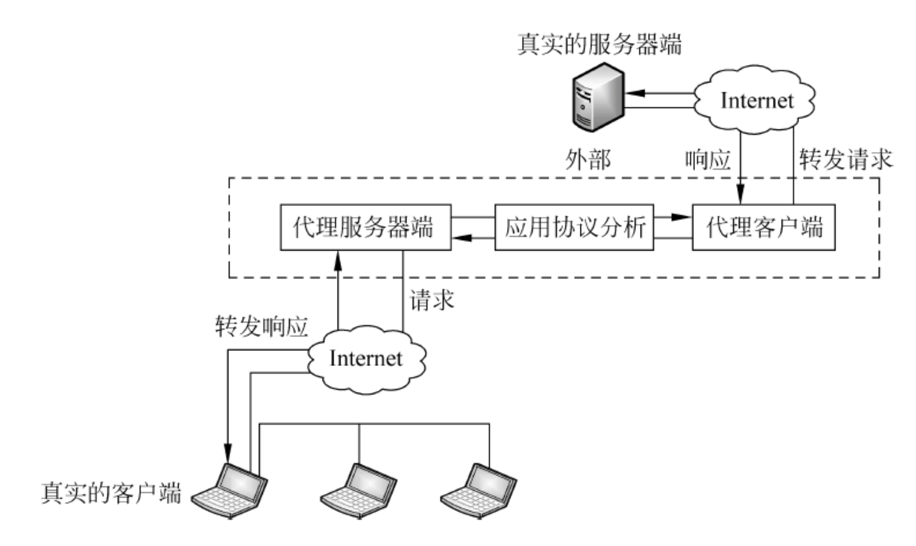


图 10.6 代理防火墙的工作方式

另一种类型的代理技术称为电路层网关(Circuit Gateway)。在电路层网关中,包被提交至用户应用层处理。电路层网关用来在两个通信的终点之间转换包,如图 10.7 所示。

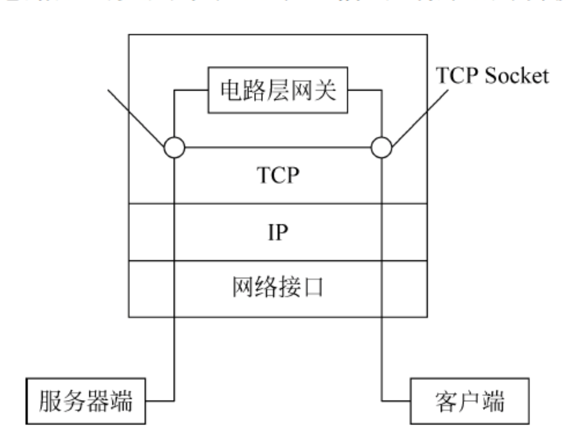


图 10.7 电路层网关

电路层网关是建立应用层网关的一个更加灵活和一般的方法。虽然它们可能包含支持某些特定 TCP/IP 应用程序的代码,但通常要受到限制。如果支持应用程序,那也很可能是 TCP/IP 应用程序。

在电路层网关中,特殊的客户机软件可能要安装,用户可能需要一个可变用户接口来相互作用或改变他们的工作习惯。代理技术的优点如下。

(1) 代理易于配置。代理因为是一个软件,所以它较过滤路由器更易配置,配置界面十分友好。如果代理实现得好,则对配置协议的要求可以低一些,从而避免了配置错误。

(2) 代理能生成各项记录。因代理工作在应用层,它检查各项数据,所以可以按一定准则,让代理生成各项日志、记录。这些日志、记录对于流量分析、安全检验是十分重要和宝贵的。当然,也可以用于计费等应用。

(3) 代理能灵活、完全地控制进出流量、内容。通过采取一定的措施,按照一定的规则,可以借助代理实现一整套的安全策略,如可以说控制"谁"和"什么",还有"时间"和"地点"。

(4) 代理能过滤数据内容。可以把一些过滤规则应用于代理,让它在高层实现过滤功能,例如,文本过滤、图像过滤(目前还未实现,但这是一个热点研究领域)、预防病毒或扫描病毒等。

(5) 代理能为用户提供透明的加密机制。用户通过代理进出数据,可以让代理完成加解密的功能,从而方便用户,确保数据的机密性。这点在虚拟专用网中特别重要。代理可以广泛地用于企业外部网中,提供较高安全性的数据通信。

(6) 代理可以方便地与其他安全手段集成。目前的安全问题解决方案很多,如认证(Authentication)、授权(Authorization)、账号(Accounting)、数据加密、安全协议(SSL)等。如果联合使用代理与这些手段,将大大增加网络安全性。这也是近期网络安全的发展方向。

代理技术的缺点如下。

(1) 代理速度较路由器慢。路由器只是简单查看 TCP/IP 报头,检查特定的几个域,不做详细分析、记录。而代理工作于应用层,要检查数据包的内容,按特定的应用协议(如 HTTP)进行审查、扫描数据包内容,并进行代理(转发请求或响应),故其速度较慢。

(2) 代理对用户不透明。许多代理要求客户端做相应改动或安装定制客户端软件,这给用户增加了不透明度。为庞大的互异网络的每一台内部主机安装和配置特定的应用程序既耗费时间,又容易出错,原因是硬件平台和操作系统都存在差异。

(3) 对于每项服务代理可能要求不同的服务器。可能需要为每项协议设置一台不同的代理服务器,因为代理服务器不得不理解协议以便判断什么是允许的和不允许的,并且还装扮成一个对真实服务器来说是客户、对代理客户来说是服务器的角色。挑选、安装和配置所有这些不同的服务器也可能是一项较大的工作。

(4) 代理服务通常要求对客户、过程之一或两者进行限制。除了一些为代理而设的服务外,代理服务器要求对客户与/或过程进行限制,每一种限制都有不足之处,人们无法经常按他们自己的步骤使用快捷可用的工作。由于这些限制,代理应用就不能像非代理应用运行得那样好,它们往往可能曲解协议的说明,并且一些客户和服务器比其他的要缺少一些灵活性。

(5) 代理服务不能保证免受所有协议弱点的限制。作为一个安全问题的解决方法,代理取决于对协议中哪些是安全操作的判断能力。每个应用层协议,都或多或少存在一些安全问题,对于一台代理服务器来说,要彻底避免这些安全隐患几乎是不可能的,除非关掉这些服务。

代理取决于在客户端和真实服务器端之间插入代理服务器的能力,这要求两者之间交流的相对直接性,而且有些服务的代理是相当复杂的。

(6) 代理不能改进底层协议的安全性。因为代理工作在 TCP/IP 之上,属于应用层,所以它不能改善底层通信协议的能力。如 IP 欺骗、SYN 泛滥、伪造 ICMP 消息和一些拒绝服务的攻击。而这些方面,对于一个网络的健壮性是相当重要的。试想,如果网络都不能正常运行,代理服务从何谈起呢?

10.5.2　防火墙的体系结构

每个网络在商业模式上都是独特的，防火墙也应该按照公司的特别要求而制作。当设计一个防火墙方案时，必须考虑很多的因素，包括费用、培训、安全、技术和完成所需要的时间。

防火墙的 4 种基本体系结构如下。

（1）屏蔽路由器。

（2）双穴主机网关。

（3）屏蔽主机网关。

（4）被屏蔽子网（非军事区 DMZ，Demilitarized Zone 结构模式）。

1）包过滤路由器防火墙

包过滤路由器是一种便宜、简单、常见的防火墙。包过滤路由器在网络之间完成数据包转发的普通路由功能，并利用包过滤规则允许或拒绝数据包。其结构如图 10.8 所示。

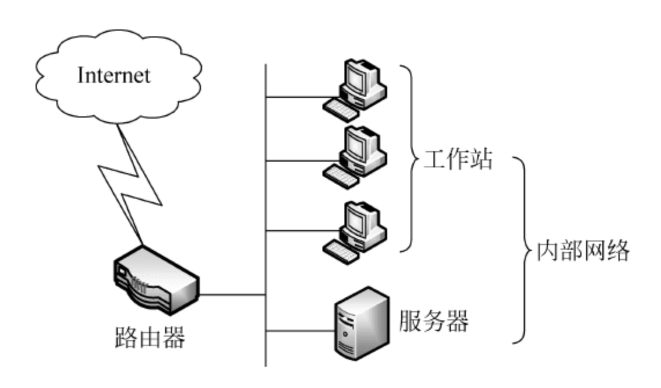


图 10.8　包过滤路由器防火墙

尽管这种防火墙系统有价格低和易于使用的优点，但同时也有缺点，如配置不当的路由器可能受到攻击，以及利用攻击包裹在允许的服务和系统内进行攻击等。由于允许在内部和外部系统之间直接交换数据包，因此攻击面可能会扩展到所有主机和路由器所允许的全部服务上。这就意味着可以从外部网上直接访问的主机要支持复杂的用户认证，并且网络管理员要不断地检查网络以确定网络是否受到攻击。另外，如果有一个包过滤路由器被渗透，则内部网上的所有系统都可能会受到损害。

2）屏蔽主机防火墙

屏蔽主机防火墙系统采用了包过滤路由器和堡垒主机，组成如图 10.9 所示。这个防火墙系统提供的安全等级比包过滤路由器要高，因为它实现了网络层安全（包过滤）和应用层安全（代理服务）。所以入侵者在破坏内部网络的安全性之前，必须首先渗透两种不同的安全系统。

对于这种防火墙系统，堡垒主机配置在内部网上，而包过滤路由器则放置在内部网和外部网之间。在路由器上进行规则配置，使得外部系统只能访问堡垒主机，去往内部系统上其他主机的信息全部被阻塞。由于内部主机与堡垒主机处于同一个网络，内部系统是否允许直接访问外部网，或者是要求使用堡垒主机上的代理服务来访问外部网由机构的安全策略来决定。对路由器的过滤规则进行配置，使得其只接收来自堡垒主机的内部数据包，就可以强制内部用户使用代理服务。

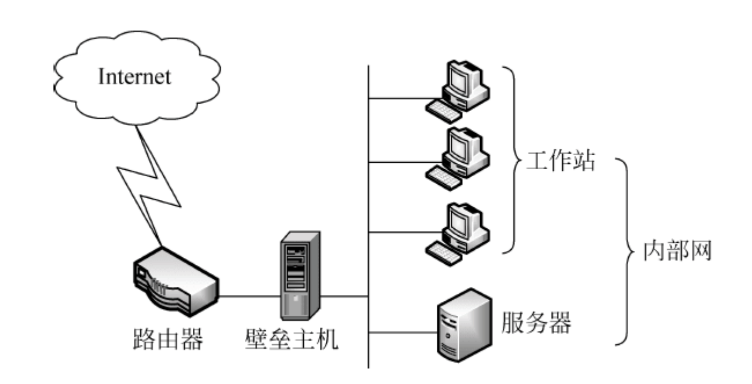


图 10.9 屏蔽主机防火墙(单宿堡垒主机)

这种防火墙系统的优点之一是提供公开的信息服务的服务器,如 Web、FTP 等,可以放置在由包过滤路由器和堡垒主机共用的网段上。如果要求有特别高的安全特性,可以让堡垒主机运行代理服务,使得内部和外部用户在与信息服务器通信之前,必须先访问堡垒主机。如果较低的安全等级已经足够,则将路由器配置成让外部用户直接去访问公共的信息服务器。

用双宿堡垒主机甚至可以构造更加安全的防火墙系统,如图 10.10 所示。双宿堡垒主机有两个网络接口,但是主机在两个端口之间直接转发信息的功能(其能旁路代理服务)被关掉了。这种物理结构强行将让所有去往内部网的信息经过堡垒主机,并且在外部用户被授予直接访问信息服务器的权利时,提供附加的安全性。

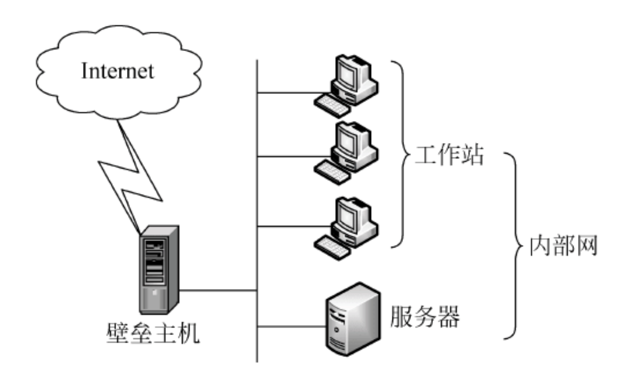


图 10.10 屏蔽主机防火墙(双宿堡垒主机)

由于堡垒主机是唯一能从外部网上直接访问的内部系统,因此有可能受到攻击的主机就只有堡垒主机本身。但是,如果允许用户注册到堡垒主机,那么整个内部网上的主机都会受到攻击的威胁。这是因为,对于入侵者来说,如果允许注册,破坏堡垒主机相对比较容易。牢固可靠、避免被渗透和不允许用户注册对堡垒主机来说是至关重要的。

3) 屏蔽子网防火墙

屏蔽子网防火墙系统采用了两台包过滤路由器和一台堡垒主机,如图 10.11 所示。这个防火墙系统建立的是最安全的防火墙系统,因为在定义了“非军事区”(DMZ)网络后,它支持网络层和应用层安全功能。网络管理员将堡垒主机、信息服务器、Modem 组,以及其他公用服务器放在 DMZ 网络中。DMZ 网络很小,处于外部网和内部网之间。在一般情况下将 DMZ 配置成使用外部网和内部网系统能够访问 DMZ 网络上数目有限的系统,而通过 DMZ 网络直接进行信息传输是严格禁止的。

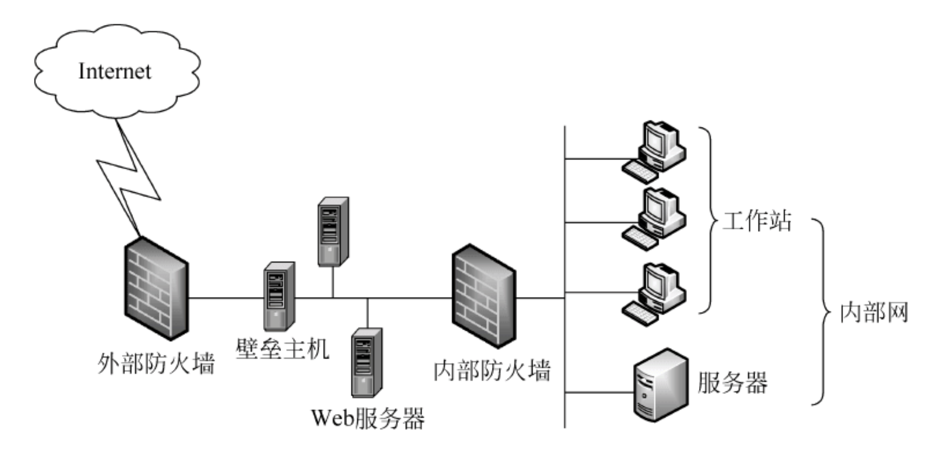


图 10.11 屏蔽子网防火墙

对于进来的信息，外面的这台路由器用于防范通常的外部攻击(如源地址欺骗和源路由攻击)，并管理外部网到 DMZ 网络的访问。它只允许外部系统访问堡垒主机(还可能有信息服务器)。里面的这台路由器提供第二层防御，只接收源于堡垒主机的数据包，负责管理 DMZ 到内部网的访问。

对于去往外部网的数据包，里面的路由器管理内部网到 DMZ 网络的访问，它允许内部系统只访问堡垒主机(还可能有信息服务器)。外面的路由器上的过滤规则要求使用代理服务(只接收来自堡垒主机的去往外部网络的数据包)。

部署屏蔽子网防火墙系统有如下好处。

(1) 入侵者必须突破 3 个不同的设备才能侵袭内部网：外部路由器、堡垒主机，还有内部路由器。由于外部路由器只能向外部网通告 DMZ 网络的存在，外部网上的系统不需要有路由器与内部网相对。这样网络管理员就可以保证内部网是“不可见”的，并且只有在 DMZ 网络上选定的系统才对外部网开放(通过路由表和 DNS 信息交换)。由于内部路由器只向内部网通告 DMZ 网络的存在，内部网上的系统不能直接通往外部网，这样就保证了内部网上的用户必须通过驻留在堡垒主机上的代理服务才能访问外部网。

(2) 包过滤路由器直接将数据引向 DMZ 网络上所指定的系统，消除了堡垒主机双宿的必要。内部路由器在作为内部网和外部网之间最后的防火墙系统时，能够支持比双宿堡垒主机更大的数据包吞吐量。

由于 DMZ 网络是一个与内部网不同的网络，NAT(网络地址变换)可以安装在堡垒主机上，从而避免在内部网上重新编址或重新划分子网。

10.6 网络攻击与入侵检测技术

10.6.1 网络攻击方法

1. 端口扫描

一个端口就是一个潜在的通信通道，也就是一个入侵通道。对目的计算机进行端口扫描能得到许多有用的信息。进行扫描的方法很多，可以是手动进行扫描，也可以用端口扫描

软件进行扫描。手动进行扫描需要熟悉各种命令，对命令执行后的输出进行分析。用扫描软件进行扫描时，许多扫描软件都有分析数据的功能。通过端口扫描，可以得到许多有用的信息，从而发现系统的安全漏洞。

2. 口令破解

通过破解获得系统管理员口令，进而掌握服务器的控制权是黑客的一个重要手段。破解获得管理员口令的方法有很多，下面是3种最为常见的方法。

(1) 猜解简单口令：很多人使用自己或家人的生日、电话号码、房间号码、简单数字或者身份证号码中的几位；也有的人使用自己、孩子、配偶或宠物的名字；还有的系统管理员使用Password，甚至不设密码，这样黑客可以很容易通过猜想得到密码。

(2) 字典攻击：如果猜解简单口令攻击失败后，黑客开始试图字典攻击，即利用程序尝试字典中的单词的每种可能组合。字典攻击可以利用重复的登录或者搜集加密的口令，并且试图同加密后的字典中的单词匹配。黑客通常利用一个英语词典或其他语言的词典。他们也使用附加的各类字典数据库，比如名字和常用的口令。

(3) 暴力猜解：同字典攻击类似，黑客尝试所有可能的字符组合方式。一个由4个小写字母组成的口令可以在几分钟内被破解，而一个较长的由大小写字母组成的口令，包括数字和标点，其可能的组合达10万亿种。如果每秒钟可以试100万种组合，可以在一个月内破解。

3. 特洛伊木马

特洛伊木马是一个包含在一个合法程序中的非法的程序。该非法程序被用户在不知情的情况下执行。其名称源于古希腊的特洛伊木马神话，传说希腊人围攻特洛伊城，久久不能得手。后来想出了一个木马计，让士兵藏匿于巨大的木马中。大部队假装撤退而将木马“丢弃”于特洛伊城，让敌人将其作为战利品拖入城内。木马内的士兵趁夜晚敌人庆祝胜利而放松警惕时从木马中爬出来，与城外的部队里应外合而攻下了特洛伊城。

一般的木马程序都包括客户端和服务器端两个程序，其中客户端是用于攻击者远程控制植入木马的机器，服务器端程序即是木马程序。攻击者要通过木马攻击用户的系统，将程序植入用户的计算机里。

目前木马入侵的主要途径还是先通过一定的方法把木马执行文件复制到被攻击者的计算机系统里，利用的途径有邮件附件、下载软件等，然后通过一定的提示故意误导被攻击者打开执行文件，比如故意谎称这个木马执行文件是朋友送的贺卡，可能在打开这个文件后确实有贺卡的画面出现，但这时可能木马已经悄悄的在后台运行了。一般的木马执行文件非常小，大部分是几千字节到几十千字节，如果把木马捆绑到其他正常文件上，用户很难发现，所以有一些网站提供的软件下载往往是捆绑了木马文件的，用户执行这些下载的文件，也同时运行了木马。

木马也可以通过Script、ActiveX及Asp.CGI交互脚本的方式植入，由于微软的浏览器在执行Script脚本时存在一些漏洞，攻击者可以利用这些漏洞传播病毒和木马，甚至直接对浏览者计算机进行文件操作等控制。前不久出现一个利用微软Script脚本漏洞对浏览者硬盘进行格式化的HTML页面。如果攻击者有办法把木马执行文件下载到攻击主机的一个可执行WWW目录夹里面，他可以通过编制CGI程序在攻击主机上执行木马目录。此外，木马还可以利用系统的一些漏洞进行植入，如微软著名的US服务器溢出漏洞，通过一个

IIS Hack 攻击程序即可使 IIS 服务器崩溃，并且同时攻击服务器，执行远程木马文件。

当服务器端程序在被感染的机器上成功运行以后，攻击者就可以使用客户端与服务器端建立连接，并进一步控制被感染的机器。在客户端和服务器端通信协议的选择上，绝大多数木马使用的是 TCP/IP，但是也有一些木马由于特殊的原因，使用 UDP 进行通信。当服务器端在被感染机器上运行以后，它一方面尽量把自己隐藏在计算机的某个角落里面，以防被用户发现；另一方面，监听某个特定的端口，等待客户端取得连接；另外为了下次重启计算机时仍然能正常工作，木马程序一般会通过修改注册表或者其他的方法让自己成为自启动程序。

4. 缓冲区溢出攻击

1）缓冲区溢出攻击的原理

缓冲区是程序运行时机器内存中的一个连续块，它保存了给定类型的数据，随着动态分配变量会出现问题。大多数时候为了不占用太多的内存，一个有动态分配变量的程序在程序运行时才决定给它们分配多少内存。这样下去，如果说要给程序在动态分配缓冲区放入超长的数据，它就会溢出了。

缓冲区溢出是非常普遍和危险的漏洞，在各种操作系统、应用软件中广泛存在。产生缓冲区溢出的根本原因在于，将一个超过缓冲区长度的字符串复制到缓冲区，就会溢出。造成两种后果，一是过长的字符串覆盖了相邻的存储单元，引起程序运行失败，严重的可引起死机、系统重新启动等；二是利用这种漏洞可以执行任意指令，甚至可以取得系统特权，使用一类精心编写的程序，可以很轻易地取得系统的超级用户权限。

缓冲区溢出攻击指的是一种系统攻击的手段，通过往程序的缓冲区写入超出其长度的内容，造成缓冲区的溢出，从而破坏程序的堆栈，使程序转而执行其他指令，以达到攻击的目的。据统计，通过缓冲区溢出进行的攻击占所有系统攻击总数的 80%以上。

缓冲区是内存中存放数据的地方。在程序试图将数据放到机器内存中的某一个位置时，因为没有足够的空间就会发生缓冲区溢出。而人为的溢出则是有一定企图的，黑客写一个超过缓冲区长度的字符串，然后植入缓冲区。缓冲区溢出成为远程攻击的主要手段，其原因在于缓冲区溢出漏洞给予了黑客所想要的一切：植入并且执行攻击代码。被植入的攻击代码以一定的权限运行有缓冲区溢出漏洞的程序，从而得到被攻击主机的控制权。大多数造成缓冲区溢出的原因是程序中没有仔细检查用户输入的参数。

2）缓冲区溢出攻击的方法

缓冲区溢出攻击的目的在于扰乱具有某些特权的运行程序的运行。这样可以让黑客取得程序的控制权，如果该程序具有足够的权限，那么整个主机就被控制了。一般而言，在 UNIX 中黑客攻击 Root 程序，然后执行类似 exec(sh)的执行代码来获得 Root 的 Shell。为了达到这个目的，黑客必须在程序的地址空间里安排适当的代码或通过适当的初始化寄存器和存储器，让程序跳转到安排好的地址空间执行。

5. 拒绝服务攻击

几乎从因特网诞生以来拒绝服务攻击就伴随着因特网的发展一直存在，并也在不断地发展和升级。2000 年 2 月，Yahoo、亚马逊等受到拒绝服务攻击而停止服务，近年来国内多家大型网站也受到攻击甚至有的还影响了 Web 服务的提供。人们每年都会见到很多大网站受到拒绝服务攻击。拒绝服务即 Denial of Service，简称 DoS，由于它的不易觉察性和简

易性，因而一直是网络安全的重大隐患。它是一种技术含量低、攻击效果明显的攻击方法，受到攻击时，服务器在长时间内不能提供服务，使得合法用户不能得到服务，特别是分布式拒绝服务(DDoS)，它的效果很明显，并且难以找到真正的攻击源，很难找到行之有效的解决方法。

拒绝服务攻击是一种广泛的系统漏洞，黑客们正热衷于对它的研究，而无数的网络用户将成为这种攻击的受害者。它是一种简单的破坏性攻击，通常黑客利用 TCP/IP 中的某种漏洞，或者系统存在的某些漏洞，对目标系统发起大规模的攻击，使攻击目标失去工作能力，使系统不可访问，因而合法用户不能及时得到应得的服务或系统资源，如 CPU 处理时间与网络带宽等。它最本质的特征是延长正常的应用服务的等待时间。

根据 TCP/IP 的原理，当客户端要和服务器端进行通信时，会经过请求/确认的方式进行联系，如用户登录服务器时，首先是用户传送信息要求服务器确认，服务器端给予响应回复客户端请求，当被确认后，客户端才能正式和服务器端交流信息。在拒绝服务攻击情况下，黑客凭借虚假地址向服务器提交连接请求，当然服务器回复信息时就送到这个虚假地址，但是服务器回传时却无法找到这个地址，根据 TCP/IP 连接原理，此时服务器会进行等待，达到超时设置时才会断开这个连接。如果攻击者传送多个这样的请求或利用多个站点同时传送这样的请求，那么服务器就会等待更长时间，这个过程周而复始，最终会导致服务器资源用尽，网络带宽用完，正常的服务请求不能被服务器处理及回复而形成服务器的拒绝服务。拒绝服务并不是服务器不接受服务，而是服务器太忙，不能及时地响应请求，相对于客户来说就认为是服务器拒绝给予服务，严重时会造成服务器死机，甚至导致整个网络瘫痪。

拒绝服务攻击的目的不在于闯入一个站点或更改数据，而在于使站点无法服务于合法的请求。入侵者并不单纯为了进行拒绝服务而入侵，拒绝服务往往是为了完成其他入侵的必需的前提。例如，在目的主机上放置了木马等恶意程序，需要让目标主机重启；为了完成 IP 欺骗，而使被冒充的主机瘫痪；在正式入侵之前，使目的主机的日志系统不能正常工作；还有可能是出于政治或经济上的目的而发动的拒绝服务。

网络对拒绝服务攻击的抵抗力很有限，黑客可以阻止合法的用户使用网络和服务。常见的针对网络的拒绝服务攻击方式有以下几种。

1）死亡之 Ping

死亡之 Ping 是最简单的攻击方式，攻击者通过 Ping 命令向被攻击者发送大量超大字节的 ICMP 报文来攻击。在许多操作系统中 ICMP 数据包默认的大小为 64KB，当然用户也可以利用这个命令发送大字节的数据包，但是发送的包超过 65 535B 就会导致服务器重组包时造成内存分配错误，发生缓冲区溢出，严重时会使服务器崩溃而拒绝服务。如命令：

```
Ping -165535 <目的主机的 IP 地址>
```

对这类攻击的防范比较容易，可以安装防火墙拒绝对 ICMP 报文的响应，就会把这样的数据包拦住，另外也可在操作系统里修改设置，如在 Windows 2000 中在 IP 安全策略中设定把 ICMP 的包过滤掉就可以解决。

2）SYN Flood 攻击

在 TCP/IP 协议标准中，要完成一个 TCP 连接需要客户端和服务器端三次握手。首先

客户端发一个有 SYN 标志的包给服务器端请求服务，然后服务器端返回一个 SYN+1 的 ACK 响应包，客户端收到后再发一个确认包给服务器端。这时，客户端与服务器端建立连接成功，这样就为以后的通信做好了准备工作。但在进行攻击时，攻击者利用假冒的 IP 地址只发伪造的包而不接收响应，从而让服务器产生大量的“半开连接”，由于每台包服务器有一定的等待响应时间，而且当一定时间没有收到响应时，还会多次重发。因此服务器端在重发与等待过程中形成大量的半开连接。就这样攻击者可通过多台计算机发送大量的虚假 IP 源地址的 SYN 数据包，造成服务器 CPU 的占用过度，当达到一定量时，就形成了拒绝服务。同时由于合法用户的请求大多被攻击包淹没，即便服务器此时没有死机，也无力再响应合法用户的请求了。

用户可以使用 Netstat 命令来检查连接线路的目前状况，看看是否有 SYN Flood 攻击。只要线路处于 SYN-Received 状态下，则表明系统可能正遭到攻击。

实施这种攻击的黑客无法取得系统中的任何访问权。但是对于大多数的 TCP/IP 协议栈来说，处于 SYN-Received 状态的连接数量非常有限。对于这种攻击，可以减少服务器重发包的次数和等待的时间，由于 SYN Flood 入侵的效果取决于服务器上保持的 SYN 半连接数据请求量，这个值等于 SYN 入侵的频度×SYN 超时时间，所以通过缩短从接收 SYN 报文到确定这个报文无效并丢弃该连接的时间，可以成倍地降低服务器的负荷。在 Windows 2000 中可修改注册表的相应设置来防范这种攻击。

3）Land 攻击

Land 攻击是利用 TCP/IP 的漏洞，发送大量的源地址与目的地址相同的数据包，从而造成服务器解析 Land 包时占用大量的处理资源，当收到的包达到一定程度时，就会形成拒绝服务攻击。在 Land 攻击中，一个特别制定的 SYN 包中的源地址和目的地址都被设置成某一个服务器地址，这时将导致接收服务器向它自己的地址发送 SYN-ACK 消息，结果这个地址发回 ACK 消息并创建一个空链接，每一个这样的链接都将保留直到超时。

对这类攻击的防范可通过防火墙解决，将那些在外部接口上入站的含有源地址是内部地址的数据包丢掉不做处理，这就是简单的防火墙的包过滤功能，另外要打上最新的补丁。

4）Smurf 攻击

Smurf 攻击结合了 IP 欺骗和 ICMP 回复方法，使大量网络传输充斥目标系统，导致目标系统拒绝为正常系统服务。在这种攻击中，入侵者从远程的网络地址发送 IP 广播地址的 ICMP echo 报文，当它收到此类数据包再转发出去时，该网络上所有的计算机将收到此数据包再传回 ICMP echo 响应报文，这将使网络拥塞不通。而且，入侵者在传送 ICMP echo 报文时，使用目的主机的 IP 地址作为源地址，这样，整个网络的所有计算机在回复 ICMP echo 报文时所传回的目的地址是目的主机，造成目的主机的网络因涌入大量的 ICMP 报文而无法使用。这种方法的可怕之处在于利用体积很小的网络数据包在很短时间里创造出大量的数据包流。

对这种攻击，可以采取一些隔离设备，使之不能进行广播。如把网络划分成多个 VLAN，每个 VLAN 中的广播不会传播到其他广播域中，以此来解决这样的攻击。解决本问题的主要思路就是防止计算机以 IP 广播请求做出响应。

5）Teardrop 攻击

数据链路层的最大传输单元 MTU 限制了数据帧的最大长度，不同的网络类型都有一

个上限值。可以用 Netstat-i 命令查看这个值,以太网的 MTU 是 1 500。如果 IP 层有数据包要传,而且数据包的长度超过了 MTU,那么 IP 层就要对数据包进行分片操作,使每一片的长度都小于或等于 MTU。每个 IP 分片都各自发送,到达目的主机后在 IP 层重组,IP 数据包头部中的数据能够正确完成分片的重组。若在 IP 分组中指定一个非法的偏移值,将可能造成某些协议软件出现缓冲区覆盖,导致系统崩溃。

解决的办法就是对系统打上最新的补丁,禁止防火墙的重组碎片功能,可以在 Windows 2000 中,自定义 IP 安全策略并设置碎片检查。

6) UKP Flood 攻击

在 UKP Flood 攻击中,攻击者可发送大量伪造源 IP 地址的小字节 UDP 包。由于 UDP 是无连接性的,所以只要打开了一个 UDP 的产品并提供相关服务,那么就可针对相关的服务进行攻击,如 QQ 就是基于 UDP 的,网上有些工具就可以发送大量的包对目标进行攻击,从而让在线 QQ 因资源不足被迫下线,如果是对其他的服务进行攻击,严重时还可能使服务器死机。

6. 网络监听

根据 IEEE 的描述,局域网技术是"把分散在一个建筑物或相邻几个建筑物中的计算机、终端、大容量存储器的外围设备、控制器、显示器,以及为连接其他网络而使用的网络连接器等相互连接起来,以很高的速度进行通信的手段"。

局域网具有设备共享、信息共享,可进行高速数据通信和多媒体信息通信,分布式处理,具有较高的兼容性和安全性等基本功能和特点。目前,局域网主要用于办公室自动化和校园教学及管理,根据具体情况采用总线型、环形、树形及星形等不同拓扑结构。

1) 网络监听

网络监听技术本来是提供给网络安全管理人员进行管理的工具,可以用来监视网络状态、数据流动情况以及网络上传输的信息等。当信息以明文的形式在网络上传输时,使用监听技术进行攻击并不是一件难事,只要将网络接口设置成监听模式,便可以源源不断地将网上传输的信息截获。网络监听可以在网上的任何一个位置实施,如局域网中的一台主机、网关上或远程网的调制解调器之间等。

2) 在局域网实现监听的基本原理

对于目前很流行的以太网协议,其工作方式是:将要发送的数据包发往连接在一起的所有主机,包中包含着应该接收数据包主机的正确地址,只有与数据包中目的地址一致的那台主机才能接收。但是,当主机工作在监听模式下,无论数据包中的目的地址是什么,主机都将接收(当然只能监听经过自己网络接口的那些包)。

10.6.2 入侵检测系统概述

什么是入侵检测?ICSA 入侵检测系统论坛的定义是:通过从计算机网络或计算机系统中的若干关键点收集信息并对其进行分析,以发现网络或系统中是否有违反安全策略的行为和遭到袭击的迹象。

入侵检测系统(Intrusion Detection System,IDS)主要通过以下几种活动来完成任务:监视、分析用户及系统活动;对系统配置和弱点进行审计;识别与已知的攻击模式匹配的

活动；对异常活动模式进行统计分析；评估重要系统和数据文件的完整性；对操作系统进行审计跟踪管理，并识别用户违反安全策略的行为。除此之外，有的入侵检测系统还能够自动安装厂商提供的安全补丁软件，并自动记录有关入侵者的信息。

入侵检测是对防火墙的合理补充，帮助系统对付网络攻击，扩展了系统管理员的安全管理能力（包括安全审计、监视、进攻识别和响应），提高了信息安全基础结构的完整性。它从计算机网络系统中的若干关键点收集信息，并分析这些信息，看看网络中是否有违反安全策略的行为和遭到袭击的迹象。入侵检测被认为是防火墙之后的第二道安全闸门，在不影响网络性能的情况下能对网络进行监测，从而提供对内部攻击、外部攻击和误操作的实时保护。这些都通过它执行以下任务来实现。

（1）监视、分析用户及系统活动。

（2）系统构造和弱点的审计。

（3）识别反映已知进攻的活动模式并向相关人士报警。

（4）异常行为模式的统计分析。

（5）评估重要系统和数据文件的完整性。

（6）操作系统的审计跟踪管理，并识别用户违反安全策略的行为。

对一个成功的入侵检测系统来讲，它不但可使系统管理员时刻了解网络系统（包括程序、文件和硬件设备等）的任何变更，还能给网络安全策略的制定提供指南。更为重要的一点是，它应该管理和配置都要简单，从而使非专业人员非常容易地获得网络安全。而且，入侵检测的规模还应根据网络威胁、系统构造和安全需求的改变而改变。入侵检测系统在发现入侵后，会及时做出响应，包括切断网络连接、记录事件和报警等。

从计算机安全的目标来看，入侵是指企图破坏资源的完整性、保密性、可用性的任何行为，也指违背系统安全策略的任何事件。从入侵策略的角度看，入侵可分为企图进入、冒充合法用户、成功闯入等方面。入侵者一般称为黑客或解密高手。Anderson 把入侵者分为伪装者、违法者和秘密用户 3 类。

入侵检测是指对计算机和网络资源的恶意使用行为进行识别与响应的处理过程。它不仅能检测来自外部的入侵行为，同时也能检测内部用户的未授权活动，是一种增强系统安全的有效方法。入侵检测是从计算机网络或计算机系统中若干关键点收集信息并对其进行分析，从中发现网络或系统中是否有违反安全策略的行为和遭到攻击的迹象，同时做出响应。入侵检测的一般过程包括信息收集、信息预处理、数据检测分析和响应等，如图 10.12 所示。

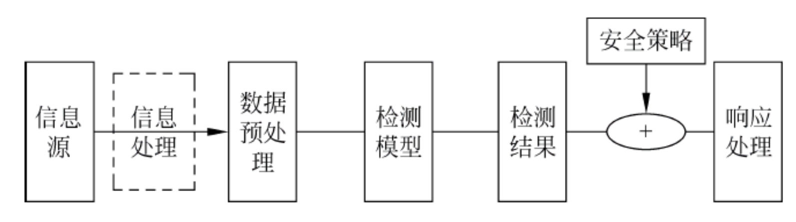


图 10.12　入侵检测的一般过程

入侵检测可分为实时入侵检测和事后入侵检测。实时入侵检测在网络连接过程中进行，系统根据用户的历史行为模型、存储在计算机中的专家知识以及神经网络模型对用户当前的操作进行判断，一旦发现入侵迹象立即断开入侵者与主机的连接，并收集证据和实施数据恢复。事后入侵检测由网络管理人员定期或不定期进行，根据计算机系统对用户操作所

做的历史审计记录判断用户是否具有入侵行为，如果有就断开连接，并记录入侵证据和进行数据恢复，但是其入侵检测的能力不如实时入侵检测。

漏洞是由软件编写不当或软件配置不当造成的。漏洞扫描是网络安全防御中的一项重要技术，其原理是采用模拟攻击的形式对目标可能存在的、已知的安全漏洞进行逐项检查，根据检测结果向系统管理员提供周密可靠的安全性分析报告，为提高网络安全的整体水平提供了重要依据。漏洞扫描也称为事前检测系统、安全性评估或者脆弱性分析。其作用是在发生网络攻击事件前，通过对整个网络进行扫描及时发现网络中存在的漏洞隐患，及时给出漏洞相应的修补方案，网络人员根据方案可以进行漏洞的修补。

漏洞检测技术通常采用两种策略，即被动式策略和主动式策略。被动式策略是基于主机的检测，对系统中不合适的设置、脆弱的口令以及其他同安全规则相抵触的对象进行检查；而主动式策略是基于网络的检测，通过执行一些脚本文件对系统进行攻击，并记录它的反应，从而发现其中的漏洞。

1. 常用的入侵检测技术

入侵检测技术可分为以下 4 种。

(1) 基于应用的监控技术，主要使用监控传感器在应用层收集信息。

(2) 基于主机的监控技术，主要使用主机传感器监控本系统的信息。

(3) 基于目标的监控技术，主要针对专有系统属性、文件属性、敏感数据等进行监控。

(4) 基于网络的监控技术，主要利用网络监控传感器监控收集的信息。

综合以上 4 种方法进行监控，其特点是提高了检测性能，但会产生非常复杂的网络安全方案。

2. 入侵检测技术的选用

在使用入侵检测技术时，应该注意以下技术特点。

(1) 信息的收集分析时间。

(2) 分析类型。

(3) 检测系统对攻击和误用的反应。

(4) 检测系统的管理和安装。

(5) 检测系统的完整性。

(6) 设置诱骗服务器。

信息的收集分析时间可分为固定时间间隔和实时收集分析两种。分析类型可分为签名分析、统计分析和完整性分析。完整性就是系统自身的安全性。设置诱骗服务器的目的就是吸引黑客的注意力，把攻击导向它，从敏感的传感器中发现攻击者的攻击位置、攻击路径和攻击实质，随后把这些信息送到一个安全的地方，供以后查用。

入侵检测作为一种积极主动的安全防护技术，提供了对内部攻击、外部攻击和误操作的实时保护，在网络系统受到危害之前拦截和响应入侵。从网络安全立体纵深、多层次防御的角度出发，入侵检测理应受到人们的高度重视，这从国外入侵检测产品市场的蓬勃发展上就可以看出。在国内，随着上网的关键部门、关键业务越来越多，迫切需要具有自主版权的入侵检测产品。但现状是入侵检测仅仅停留在研究和实验样品（缺乏升级和服务）阶段，或者是防火墙中集成较为初级的入侵检测模块。可见，入侵检测产品仍具有较大的发展空间，从技术途径来讲，除了完善常规的、传统的技术（模式识别和完整性检测）外，应重点加强统计

分析的相关技术研究。

课 后 习 题

1．术语解释

被管对象　管理进程　代理进程　管理信息库　管理工作站　防火墙　网络入侵　对称式和非对称式加密技术　数字签名　VPN　包过滤　代理服务　计算机病毒

2．SNMP 报文的格式是什么？包括哪几种报文操作？

3．简述防火墙的基本功能及其分类。

4．防火墙的体系结构是什么？

5．网络安全的策略有哪些？

6．什么是 PKI？

7．网络病毒有哪些防范技术与方法？

8．包过滤防火墙的原理是什么？

第11章　网络系统集成及规划

本章主要介绍网络系统集成设计的一般步骤、设计原则和布线规则。通过不同规模企业网络的规划设计案例来说明一般的网络设计过程，包括桌面网络、楼宇网络及园区网络。

11.1　网络工程的概念

网络工程是从用户网络建设需求出发，充分考虑用户自身特点和行业特征，利用当前主流网络技术和网络产品，设计网络构建的解决方案，并依此方案进行整个网络建设的过程。整个网络建设的过程可以称为PDIOO，即规划（Planning）、设计（Designing）、实施（Implementing）、运行（Operating）以及优化（Optimizing），本章主要集中在规划与设计部分。

11.1.1　网络工程规划

在网络建设开始之前，网络规划首先需要进行需求分析工作，根据用户对所期望建设网络的描述以及充分的交流，找出影响网络设计的关键信息。需求分析采取自顶向下的分析方法，了解用户所在行业的特征，用户个性化的特点，这些都有助于确定用户网络建设的目标并归档，这些目标依赖于每个特定用户的组织结构和环境特征，但这些目标通常都包括以下几个内容。

（1）网络的功能：构建的网络必须能够正确工作，即能够支撑用户完成他们的工作；网络必须以合适的速率和可靠性提供用户至用户以及用户与应用的连接。

（2）网络的扩展：网络必须能够成长，即原来的设计能够在不做大刀阔斧修改的情况下，根据需要在网络规模和应用上有所增长。

（3）网络的适应能力：网络在设计上必须着眼技术的发展趋势，设计的内容中不能包含那些以后可能称为网络引入新技术的障碍的部分。

（4）网络的管理：网络在设计上应该实现对网络的监控和管理，以保证运行的稳定性。

通过对以上内容的确定，论证网络构建的科学性与正确性，提出网络建设的方案。方案往往不止一个，各自实施可能的结果也不尽相同，这时可以由用户通过决策选择出最佳的方案。

11.1.2　网络工程设计

网络设计是根据网络规划及总体方案，对网络体系结构、子网划分、逻辑网络组成及网

络技术和设备选型进行工程化设计的过程，并产生具体的解决方案。这部分主要包括网络设计原则、通信子网设计、资源子网设计、设备选型、网络应用平台设计及网络安全设计。

1. 网络设计原则

网络设计要能够保证方案的切实可行，并且能够最大限度地保护用户的已有投资，在具体内容的设计上一定要遵从以下原则。

(1) 实用性：IT 技术的高速发展使得硬件(包括计算机、服务器、网络设备、网络介质)在价格和性能方面向着两个相反的方向变化。所以，应该进行合理投入并充分发挥网络的所有功能，具体设计中应做到“够用就好”，要让有限的投资发挥出最大的效益。

(2) 开放性：网络设计应该符合 OSI 参考模型，即层次化和标准化，使得不同系统间有良好的互操作性。遵循国际、行业相关的标准，采用开放的技术、开放的体系结构、开放的系统组件和开放的用户接口。

(3) 可靠性：网络的设计要能够保证之后网络运行时的稳定可靠，具有较高的平均无故障时间和较低的平均无故障率，提供容错设计，支持故障检测和恢复。

(4) 安全性：网络在设计上要能够保证在多个层次上提供安全机制，保证用户的业务、数据等的安全，对于一些特定行业尤为重要。

(5) 先进性：网络设计中采用先进的设计思想、先进的网络技术、先进的网络产品，以延长用户网络的寿命。当然，先进性与实用性往往会有矛盾，所以在两者间进行平衡将是体现网络设计性价比的重要因素。

(6) 易用性：整个网络系统必须易于管理、变化和使用，为各种应用提供良好的用户接口，并屏蔽技术细节，使得用户使用网络达到透明化。

(7) 可扩展性：网络设计既能够保证用户今后网络规模上的增长，同时又能够保证新应用引入时所需的宽带，能够在规模和性能两个方面进行扩展。

2. 通信子网设计

由于这里建设的网络所服务的对象是诸如企业之类的用户，所以这里的通信子网不是电信行业提供的广域网，而是用户环境下大规模的园区级网络体系结构及其骨干部分。

通信子网在体系结构上使用的是层次式和模块化的模型，这里的层次并不是一台网络设备内功能的层次(OSI 层次)划分，而是通信子网内网络设备间的层次划分，划分的依据根据设备功能使用的重心而决定，比如一台三层交换机，在设计上究竟是更多地利用它的两层交换功能，还是更多地利用它的三层交换功能，这可能会决定它处在模型的哪个层次上。

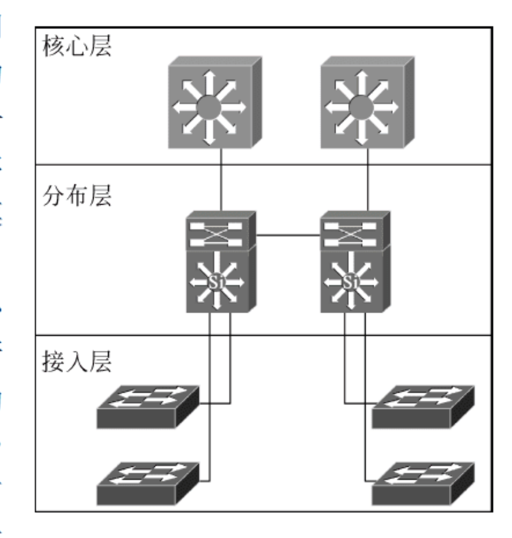


图 11.1 层次型模型

一个层次式的通信子网包括 3 个层次，即核心层、分布层及接入层，属于不同层次的设备职责有所不同，如图 11.1 所示。当然，根据通信子网规模的差别，可以将多个层次需要完成的功能部署到一台网络设备完成，比如，在规模较小的园区网络里，可以用一台高性能三层交换机同时完成核心层和分布层的功能。此外，核心层可以根据需要，选择不同的

交换技术，如图 11.2 所示，可以分别使用 ATM 交换技术，或者以太网交换技术。

核心层是一个快速交换主干，应该尽可能地以最快的速率交换数据包。核心层不应该执行任何对包的操作（比如用 ACL 进行过滤），这样会降低包交换的速率。核心层一般用来连接建筑群和服务器群，承担网络上 80%左右的流量，是网络的主干。连接键族群的主干网一般以光缆作传输介质，典型的主干网技术主要有千兆以太网、100Base-FX、ATM 和 FDDI 等。从易用性、先进性和可扩展性的角度考虑，采用千兆以太网是目前最普遍的做法。核心层的功能是由核心交换机（或路由器）来完成的。若考虑网络的容错能力，在条件允许下可以采用双核心结构，即核心层有冗余设备，并且核心层与分布层在连接上也有冗余链路。

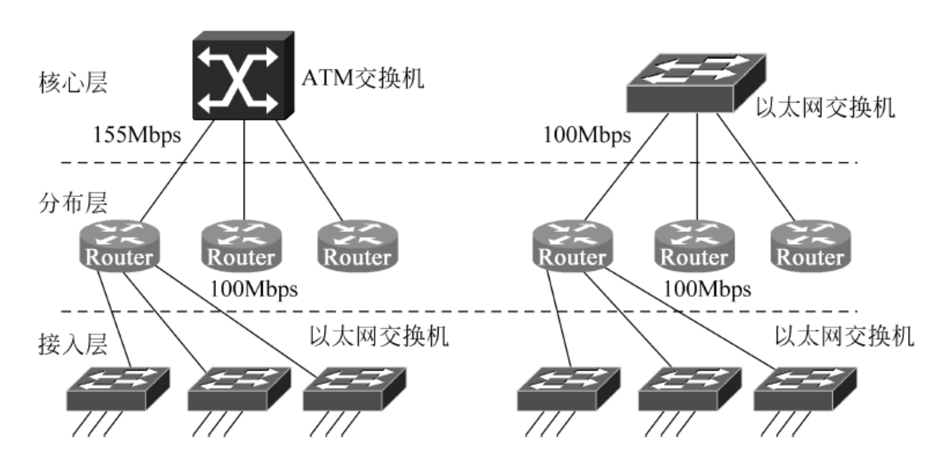


图 11.2 层次型模型实施方案

分布层用来分隔接入层和核心层，并且用来区分和定义核心层的位置，如图 11.3 所示。该层用来划分逻辑子网的边界，执行对数据包的操作，主要包括地址和区域的聚合，提供部门或工作组的访问，广播域/组播域的边界划分，VLAN 间路由，网络介质的转换，提供安全机制。分布层可以是不同路由协议进行相互进入的节点，也可以是动态路由协议和静态路由执行区域的边界。它同样也可以是远程站点访问该园区网络的进入点。总之，分布层提供的是基于策略的连接。

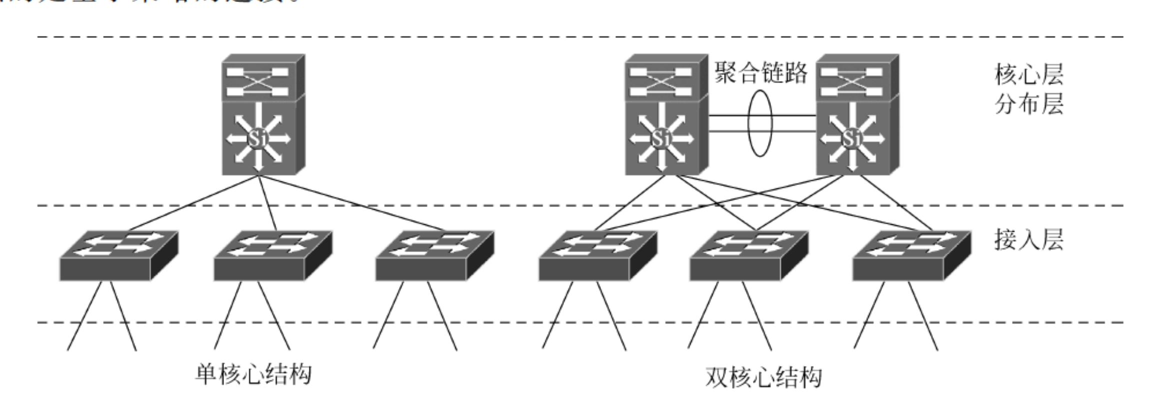


图 11.3 层次型功能的集成

接入层是本地终端用户可以接入网络的节点，该层也可以使用访问控制列表来进一步优化特殊性用户群的需求。在园区化的环境下，接入层的功能主要是利用共享和交换方式使用带宽，基于 MAC 子层的过滤以及对冲突域的微分段。

在非园区化的环境中，如个人移动用户，访问层可以利用广域网技术提供对相关网络的远程访问，可以使用的广域网技术有帧中继、ISDN、xDSL或专线。

3. 资源子网设计

1）服务器的放置

服务器系统是网络应用的核心设备，服务器在网络中的位置直接影响网络的应用性能和网络运行效率。服务一般分3类，分别为本地服务、远程服务和全局服务。本地服务是指提供服务的服务器和请求服务的终端在同一子网内或一个VLAN内，本地服务只会局限于特定的区域内，服务器的数据流量只受连接中的链路、交换机终端用户的影响。远程服务是指请求服务的用户终端和提供服务的服务器可能在地理位置上靠得很近，但它们却属于不同的子网或VLAN，数据流可能会经过主干部分，但肯定会跨越广播域的边界，这样就需要网络层的设备来提供对远程服务的访问。全局服务是指提供给网络中所有用户的公共服务，这些服务有E-mail、Internet访问、视频会议等，由于所有这些用户都会去访问全局服务，所以提供这种服务的服务器通常被统一放置在直接连接在主干上的一个独立子网，如图11.4所示。

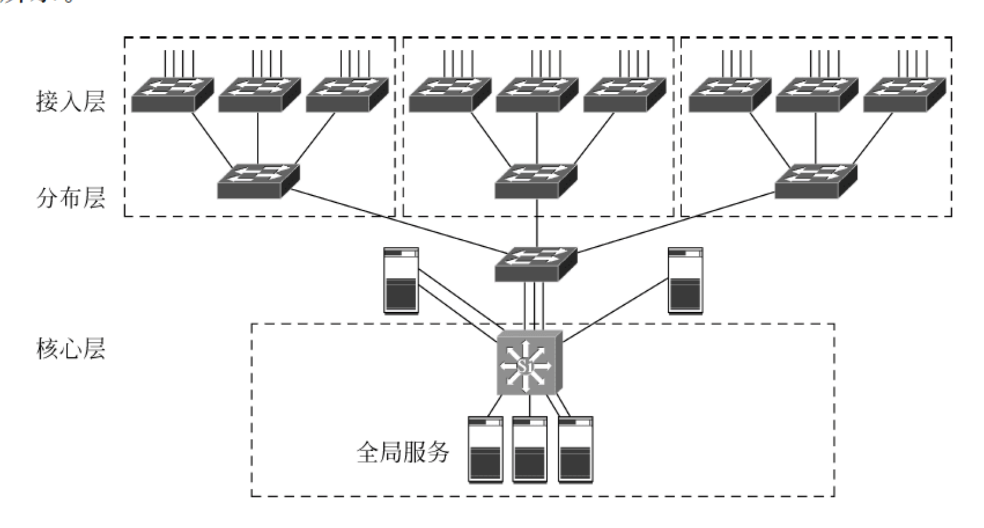


图11.4 全局服务的放置

2）服务器与核心层的连接

服务器连接到核心层设备的网络结构，可以将服务器直接连接到核心层设备，这样做可以使得服务器数据流经过的网络设备减少，减小其传输中的延迟，但这样会占用较多核心层设备上的端口；如果核心层端口密度低，并且可以接受增加一定的传输延迟，那么可以先将所有服务器设备汇聚到一台交换机，再由交换机通过高宽带链路连接到核心层设备，如图11.5所示。

4. 设备选型

本节给出了在网络设计中如何进行设备选型的原则，以及在选择层次模型中不同层次设备的要点。

1）网络设备选型原则

（1）标准化原则：所选择的设备必须基于国际标准或行业标准。因为只有基于标准的产品才符合网络开放的趋势，才有可能与其他厂商的产品互联互通，即具有良好的兼容性。

（2）权威性原则：注意厂商的选择，所使用的网络设备最好都来自同一供应商，这样在

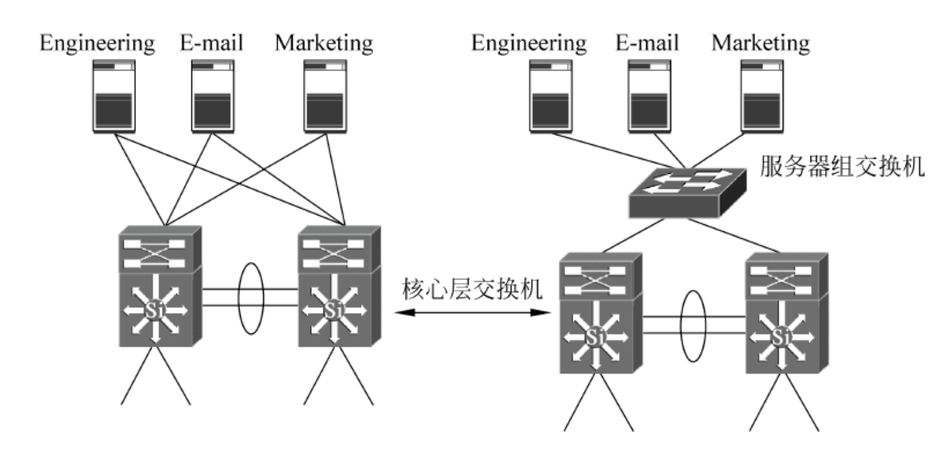


图 11.5 全局服务的子网结构

设备可互联性、协议互操作性、技术支持、价格等方面都更有优势。从这个角度来看，产品线齐全、技术认证队伍力量雄厚、产品市场占有率高的厂商是网络设备品牌的首选。其产品经过更多用户的检验，产品成熟度高，而且这些厂商出货频繁，生产量大，质保体系完备。

(3) 技术简单性原则：当对网络设备功能明确时，购买的设备并不是功能越多越好，这样只会使得今后维护的难度大大增加。

(4) 环境适应性原则：不要轻信国外某些机构的评测报告，其中不乏商业因素。而且，即使是权威机构的评测报告，也只是在特定网络环境下取得的结果，不能作为产品选型的全部依据。

(5) 实用性原则：主要是在参照整体网络设计要求的基础上，根据网络实际带宽性能需求、端口类型和端口密度选型。如果是旧网改造项目，应尽可能保留并延长用户对原有网络设备的投资，减少在资金投入方面的浪费。

(6) 可管理性原则：对于大型网络而言，这一点是至关重要的，它不仅关系到系统的性能指标，甚至关系到系统的可用性。主要考察网管系统对所选设备的监管、配置能力，以及设备可以提供的统计信息和故障检测手段，如骨干交换机必须具备端口镜像能力。这对于故障诊断，以及今后的网络规划具有特别重要的价值。

(7) 容错冗余性原则：除了在网络设计时要考虑冗余外，骨干设备的容错冗余也是必需的。所谓容错，就是设备的某一模块出现故障时，是否会影响其他模块，乃至其他设备的正常工作；是否支持热插拔；是否支持备份设备的自动切换等。所谓冗余，就是配置的设备，是否可以安装多个相同功能的模块，在工作正常的情况下实施负载分担，当其中一个出现问题时自动切换。

2) 核心层设备的选型要点

核心层设备在选择上必须支持基于 ASIC 的高性能转发能力，包括线速的 MPLS 业务处理能力、线速的 IPv6 业务处理能力、极高速的路由交换引擎；核心层设备还支持融合的网络安全特性，支持集成的防火墙模块，支持集成的 IPSec 模块，提供互联网 VPN 接入服务，支持端口镜像和远程端口镜像，能够与 IDS 联动，能够抵御网络病毒的攻击，支持路由协议报文的加密，支持以太网多种端口绑定方式，支持报文安全过滤，防止非法入侵和恶意报文攻击；产品自身的无单点故障设计，关键部分都采用冗余设计等。当然，其中一些功能只有将该设备同时作为核心层/分布层设备时才应该开启。

3）分布层设备的选型要点

应支持模块化的方式，这样可以有更好地扩展空间，并且可以很平滑地在多个异构子网间作转换，在存在用户已有网络时可以充分保护用户之前的投资；支持全分布式体系结构设计，通过主控引擎和分布式高速业务接口板上的芯片实现板内板间二三层流量的线速分布式转发，通过分布式高速业务接口板上内置的高性能 CPU 与位于主控引擎上的 CPU 协同工作，实现 ACL、流分类、QOS、组播等业务的全分布式处理；分布层设备应用有较高的交换容量，支持上百个 GE 或数十个 GE，使大型企业校园网络核心层、汇聚层网络全面升级至万兆平台成为可能。

最好支持无源背板，支持双路电源供电，支持引擎、电源、风扇的冗余，支持单板热插拔，并支持 STP/RSTP/MSTP/VRRP 等协议实现链路冗余；分布层设备应能支持遵从最小服务原则，所有可能遭受到攻击的网络在默认情况下均关闭；支持安全的 SSH 登录、基于用户安全策略的 SNMP V3、MAC＋IP＋VLAN 绑定、802.1X 认证等安全策略。支持防网络风暴攻击、防 DOS/DDOS 攻击、防扫描窥探攻击、防畸形报文攻击、防网络协议报文攻击等安全技术，支持 EAD 端点安全防御解决方案，支持内容的防火墙安全模块。

支持组播功能、802.1X、DHCH-Server、NAT、PBR、POE＋Voice Vlan、EPON 等多种业务特性，这些业务特性极大地提高了企业网络业务部署的简便性和灵活性，同时增强了对 IP 语音、视频、WLAN 的支持能力，为网络实现通信整合提供了便利。基于“ASIC＋NP”的体系结构，可以灵活地支持业务功能的不断扩展，通过多功能网络处理器模块，可以进一步支持 NAT、PBR 等多种高级业务特征，支持 CWDM 和单纤双向光模块，可以在一对光纤上承载 8 个千兆收发业务或在一根光纤上同时承载收发业务，有效节约光纤资源。支持 NetStream（网流分析）功能，通过 NetStream 与网络分析器相互配合，可帮助网络管理员轻松地获得详细的网络应用信息，使网络系统变得透明、可见。例如，查看 Web、文件传输协议（FTP）、Telnet 和其他著名的 TCP/IP 应用所占通信资源的百分比，以及用户利用网络和应用资源的详细情况，进而用于高效地规划和分配资源，并保证网络的安全运营。

支持集群管理，可以对网元进行批量配置和批量升级，实现 ACL/VLAN 的动态策略下发，同时利用网络管理平台可以实现拓扑管理、可视化图形界面、智能化性能监控、告警管理等功能，这些有助于提高网络管理人员的效率，缩短网络故障及维护扩容的时间。

4）接入层设备的选型要点

支持利用互联电缆实现多台设备的扩展，最大扩展至几百个 10/100Mbps 端口；应具有即插即用、单一 IP 管理、同步升级的优点，降低系统扩展的成本。能够通过路由热备份技术，在整个堆叠架构内实现控制平面和数据平面所有信息的冗余备份和无间断三层转发，增强堆叠架构的可靠性和性能，消除单点故障，避免业务中断。能够通过分布式链路聚合技术，实现多条上行链路的负载分担和互为备份，以提高整个网络架构的冗余性和链路资源的利用率。

除了支持传统的 802.1X 认证外，还应该检查用户终端的安全状态，通过 EAD（端点准入防御）功能，配合后台系统将防病毒、补丁修复等终端安全措施与网络接入控制、访问权限控制等网络安全措施整合为一个联动的安全体系，通过对网络接入终端的检查、隔离、修复、管理和监控，使整个网络变被动防御为主动防御，变单点防御为全面防御、变分散管理为集中策略管理，提升网络对病毒、蠕虫等新兴安全威胁的整体防御能力。

支持 PoE，通过以太网对所连接的设备（如 IP Phone、Wireless AP 等）进行远程供电，从而使得不必在使用现场为设备部署单独的电源系统，极大地减少部署终端设备布线和管理成本。通过 PoE 技术和 Voice VLAN 技术的结合可以提供完整的语音设备管理方案。支持 STP/RSTP/MSTP 两层链路保护技术，提高了链路的冗余备份，提高容错能力，保证网络的稳定运行。

支持 VRRP 虚拟路由冗余协议，与其他三层交换机构建 VRRP 备份组。构建故障时的冗余路由拓扑机构，保持通信的连续性和可靠性，有效保障网络稳定。

支持 ECMP（等价路由），通过配置多条等价路径实现上行路由的冗余备份和负载分担。

支持交流/直流双输入设计，设备既可以采用交流电源输入，也可以采用直流电源输入，二者之间热备份。支持 VCT（Virtual Cable Test）电缆检测功能，便于快速定位网络故障点。支持 DLDP（Device Link Detection Protocol，设备连接检测协议），可以监控光纤的链路状态。如果发现单向链路存在，DLDP 会根据用户配置，自动关闭或通知用户手动关闭相关端口，以防止网络问题的发生。

支持 SNMP V1/V2/V3，可支持网管平台和网管系统，支持 CLI 命令行、Web 网管、Telnet、集群管理，使设备管理更方便。通过各种开放的标准 MIB 和扩展 MIB 的支持可以提供完善的基于 SNMP 的第三方管理能力。

5. 网络应用平台设计

用户网络最终是为其提供业务上的支持，而在网络的层次结构中，直接影响用户业务方式的是网络应用平台，所以对于网络应用平台的设计非常重要，同时也是用户使用网络的直接感受。应用平台在整个网络层次的结构如图 11.6 所示。

<table>
<tr><td>信息系统</td><td>应用软件：
OA、MIS、VOD、VOIP、电子商务、远程教育、电视会议等</td><td rowspan="5">系统管理</td><td rowspan="5">安全管理</td></tr>
<tr><td>系统平台</td><td>应用服务软件：
DBMS、群件、开发工具、DNS、FTP、WWW、E-mail 等</td></tr>
<tr><td>传输平台</td><td>系统软件：
OS、TCP/IP 协议栈等</td></tr>
<tr><td rowspan="2">网络平台</td><td>LAN、WAN、网络设备、Internet 接入等</td></tr>
<tr><td>综合布线系统、传输媒体等</td></tr>
</table>

图 11.6 网络应用平台层次结构

在网络操作系统的选择上注意以下内容：网络操作系统的主要特征能够有效地支撑用户的上层应用，除了支持常规的进程、存储、设备、文件、作业管理外，是否支持 VPN、流媒体、QoS、组播、数据中心等功能；网络操作系统的生命力，即该系列操作系统维持对业务变化、增长的适应能力；支持未来网络应用所需的特征，如虚拟化技术、网络位置识别；网络操作系统的速度、性能，如支持对称多处理，支持网络负载平衡。

6. 网络安全设计

网络的普及使越来越多人的工作、学习和娱乐方式发生变化，随之安全问题也凸显出来。一方面，终端计算机面临越来越严重的安全威胁，像木马、蠕虫、病毒，以及间谍软件等各种恶意代码的泛滥，使得终端计算机成了威胁的目标，甚至成为威胁的载体，从而造成威

胁的泛滥；另一方面，随着攻击工具的普及和门槛降低，一个具有网络基本常识的人就可以对网络上的服务器和应用发起攻击，从而使这些核心服务面临着巨大的风险。因此，相对于早期的网络安全，现在的风险模型发生了根本转变，主要体现在以下3个方面。

(1) 不断变化的信任模型：如上所述，终端往往称为威胁攻击的目标和载体，在这种情况下，用户的身份被认证了，并不意味着计算机的安全状态被认证了，也不意味着用户的网络行为就能被信任，传统的基于用户名和密码的身份认证、授权、审计的信任模型已经不能满足当前的安全需求了。

(2) 不断变化的威胁模型：各种蠕虫、病毒、应用层攻击技术和E-mail、移动代码结合，形成复合攻击手段，使威胁更加危险和难以抵御。这些复合威胁直接攻击企业核心服务器和应用，给企业带来了重大损失；对网络基础设施进行DDOS攻击，造成基础设施的瘫痪；更有甚者，像电驴、BT等P2P应用和MSN、QQ等即时通信软件的普及，企业宝贵的带宽资源被无关业务的流量浪费，形成巨大的性能威胁。原来的基于网络层的“非法防伪”威胁模型已经不能适应当前的安全需要了。

(3) 不断变化的业务模型：新的网络应用带来新的应用模型，不断地会有组织构架调整和人员的变化，业务的变化带来安全策略的变化，原来的静态业务模型和安全策略已经不能满足当前的安全需求了。

因此，网络安全设计不能再使用原来的把网络和安全简单叠加的解决方案，而只有把安全融合到网络中，才能有效地解决目前的安全问题，这就是安全渗透网络。所谓安全渗透网络，不仅仅是叠加，更要实现一个整网安全的防护模型，这种模型必须具备如下几个基本特征。

(1) 信任模型要求网络必须从端点开始进行行为的管理：仅仅进行身份鉴别是不够的，必须做到对端点的行为进行识别，让行为顺从取代简单的“口令”认证，成为安全接入网络的基本条件。

(2) 威胁模型要求网络深入应用和业务内容进行保护：网络设备只工作在网络层和传输层是不够的，例如，传统的防火墙只能识别TCP80端口的流量，然后根据安全策略允许或者拒绝；但是实际上TCP80端口的流量既可能是正常的Web浏览流量，也可能是针对微软IIS服务器漏洞的一次攻击，还可以是一次基于Active X控件或JavaScript的移动代码攻击，甚至还可能是基于QQ和MSN文件传输。深入应用和业务保护，可以根据公司安全策略，对TCP 80端口的流量进行深度解析，识别出哪些是正常流量、哪些是攻击流量、哪些是网络滥用流量，从而进行精细控制，完成应用保护。

(3) 业务模型要求基础安全特征成为网络的一部分：也就是基础安全特征应该有机地渗透到网络设备中，网络可以弹性地调整和部署，以适应因为业务变化引起的安全策略的变化。

因此，在网络安全产品的选择上，应该注意以下内容。

(1) 防火墙的选择应该综合包括过滤功能、应用代理功能、状态检测功能；另外最好还包括安全身份认证、智能地址转换、丰富VPN特性、内容深度识别、流量QoS保证、强大路由能力等其他特征；同时注意防火墙产品在保证网络安全性的同时还具有很好的性能，主要要求有较高的整机吞吐量、较高的最大并发连接数及较高的每秒新建连接数。

(2) VPN产品选择上注意应具有较好的加密性能和较高的最大并发隧道数，支持的

VPN 有 L2TP、IPSec、GRE、DVPN、MPLS 和 SSL VPN；支持的加密算法包括 DES/3DES、AES、MD5、SHA-1；身份认证功能包括本地认证、RADIUS/TACACS 认证及 PKI/CA 认证。

11.2 局域网系统设计的主要内容

局域网系统设计是最常见的，也是网络工程设计的重点，因为一般的广域网连接需要借助于 ISP 或 NSP，用户只需选择相应的接入方式和业务类型即可，中间的广域连接系统无须用户考虑。

在局域网系统设计中，具体的设计内容会因不同用户需求、不同网络规模和不同应用而有所不同，特别是一些行业应用系统。基本的局域网系统一般主要包括以下几个方面。

11.2.1 网络拓扑结构设计

网络拓扑结构设计是整个设计的开始，通常人们所说的"把架子搭起来"就是这个意思。这个"架子"指的就是网络拓扑结构。具体的系统设计就是在这个"架子"上展开的。局域网和广域网拓扑结构一般有星形、总线型、树形、网状等几种，具体如何选择要根据相应的网络规模和网络应用需求而定。

11.2.2 综合布线系统设计

综合布线系统设计非常重要，设计得是否合理关系到网络应用的具体应用。综合布线系统设计主要考虑传输介质、中继传输系统、网络接口和速率的匹配等重要方面。其中最重要的是机房布线系统设计，因为所有的关键设备通常都是集中在机房中，各种布线多而杂，如何正确标识和布线是整个布线系统的关键。综合布线系统设计是根据网络用户的位置分布和传输距离进行的，当然还得考虑网络系统将来的扩展和其他布线系统，如强电系统、消防系统、电话系统等。

11.2.3 网络体系架构设计

在确定了网络拓扑结构和布线系统后，就要设计网络体系架构了，因为它关系着后面将要进行的网络设备选型和连接。在此仅以 Windows 网络操作系统平台为例进行介绍。小型局域网的域系统比较好设计，而对于大中型局域网系统则肯定不止一台域控制器，不仅如此，还可能有多个子域、多个子网、多台 DNS、DHCP、WINS 服务器。那么服务器之间的关系如何，则要好好规划，不是仅通过简单的连接就可以完成的。当然以上是假设网络操作系统是 Microsoft 公司的 Windows 网络操作系统而言的，如果是其他网络操作系统，则还需要进行具体的考虑。

11.2.4　用户系统的选择与设计

用户系统包括用户终端计算机系统和应用系统两个方面。用户终端计算机系统的选择与设计主要涉及用户计算机硬件配置、操作系统、办公系统、工具软件系统的选择等方面。应用系统并不是每个用户都需要，只是对需要的用户进行选择与设计，主要就应用系统的种类和功能模块进行选择与配置。常见的应用系统包括人事管理系统、财务管理系统、进销存管理系统、酒店管理系统、餐饮管理系统、VOD 管理系统、网站管理系统、书店管理系统等。也可能是特定开发的 ERP 系统，它将全面包括用户的所有应用系统。

11.2.5　网络设备的选型和连接

确定了网络拓扑结构、域系统和用户系统后，就可以正式选购各种网络设备了。主要的网络设备有服务器、用户计算机、网卡、交换机、路由器、防火墙等。如果有网络存储系统，则还有相应的网络存储设备，如数据存储交换机、各种媒介存储设备等。网络设备的选购最主要考虑的是网络功能和应用需求，同时兼顾设备品牌、售后服务水平和成本。

在网络设备选购中还要考虑到整个网络系统之间的接口和扩展性能问题，不要孤立地考虑某一个设备。网络接口类型的不同又直接地影响了传输介质的选择，间接地影响了总体成本。

11.2.6　数据备份与恢复系统设计

一般的企业网络系统都会有一个适合自己应用需求的数据备份和恢复系统，通常小型企业是选择 Windows 网络操作系统中的"备份"工具，而对于有特殊要求的和大中型企业用户来说，则通常是另外选择更加专业的第三方数据备份与恢复系统，甚至还可能部署复杂的网络存储系统。第三方专业的数据备份与恢复系统一般可支持异地存储、各种应用系统的数据备份与恢复、智能存储等功能。

目前可选择的第三方数据备份与恢复系统有 Veritas 的 Backup Exec 10.0、BrightStor ARCserve Backup R11、Microsoft System Center Data Protection Manager 2006、腾龙备份大师 2006、Norton Ghost 10.0、Acronis True Image 9.0 等多个。不同系统，功能不同，适合环境不同，价格和售后服务水平也不同，用户要根据自身需求来选择。

11.2.7　网络管理系统和服务器管理系统设计

网络管理系统也有高、中、低档之分，高档的网络管理系统基本都是出自大的专业公司，如 IBM Tivoli NetView、CA Unicenter、HP OpenView、AT-SNMPc 等。中档的主要是国内一些厂商开发的网络管理系统，如游龙科技的 SiteView、清华紫光比威 BitView、锐捷网络 StarView 等。而小型网络管理系统则属于网络管理工具软件，如网路岗、网络执法官等，属于共享类型的。

除了对整个网络进行管理外，现在还有一些专门针对具体服务器进行管理的系统，那就是服务器管理系统。它可以根据不同服务器的主要应用进行不同的网络监控、用户权限配置等。如网强服务器管理系统、传名内部服务器管理系统等。

11.3 楼宇网络的设计案例

楼宇网络通常包括高速接入的主设备间和阁楼层的设备间的网络，如图 11.7 所示。楼宇网络被楼层或部门分成了多个区域，主机和服务器放置在中心机房，通过物理传输介质将各区域的配线间通过中心机房联网。从各配线间引出物理传输介质到工作区的办公室。

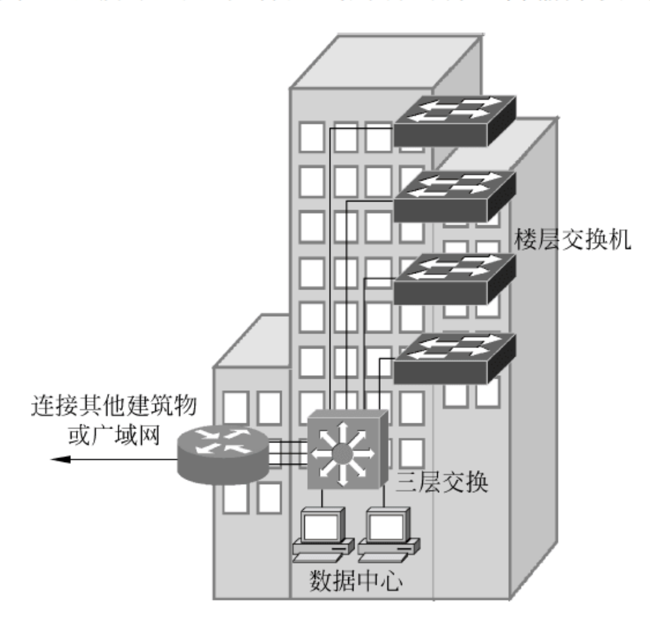


图 11.7 楼宇网络设计

某公司是一家正在快速发展的成长型公司，部门分布在公司大厦内的各个楼层，图 11.8 所示是该公司目前的网络结构图，当前整个大厦内敷设的线缆为 3 类双绞线，即局域网类型为 10Base-T；每一楼层有大约 20 台计算机，通过楼层内安装的集线器(Hub)连接起来，每个楼层的一个工作组内都安装了一台服务器，该服务器只对工作组内的用户提供服务；公司另外在第一层的机房内还部署了一定数量的服务器，为全公司提供服务；各楼层的集线器通过垂直子系统连接起来，构成整个公司的网络。

由于当前公司员工使用计算机主要是进行文档处理、数据库访问，以及收发 E-mail，每台计算机的数据流主要发往本地楼层的服务器，只有少部分的数据流发往全局服务器，即楼层 1 中的服务器，整个网络的流量模式属于 20/80 模型。另外，网络中主要采用 TCP/IP 协议栈，各计算机和服务器的 IP 地址为静态方式分配。由公司目前的网络结构可以发现，现有网络大量地使用集线器，使整个网络成为一个单独的冲突域/广播域，大量的信号冲突和广播包严重影响了网络性能，不能满足公司未来发展的需求，即 20/80 流量模式的发展。

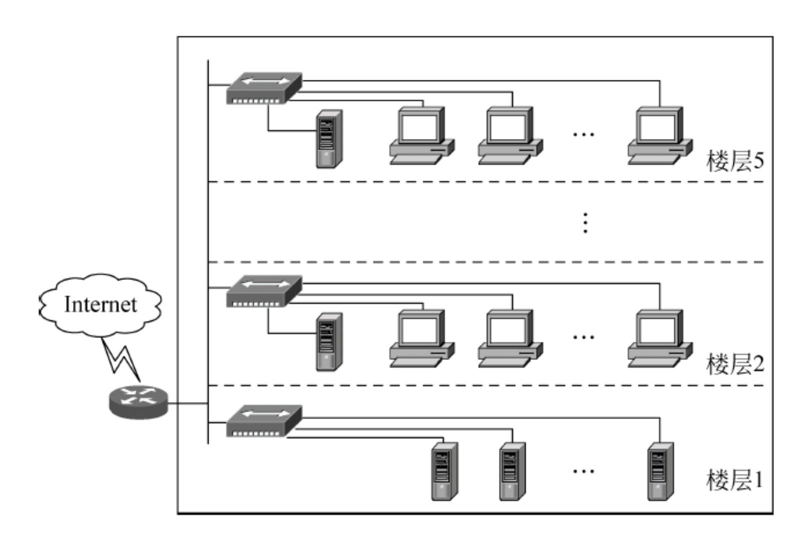


图 11.8 某公司大楼当前的网络结构

11.3.1 楼宇网络需求分析

为了适应网络 20/80 流量模式的发展以及公司业务的发展，公司管理者和网络设计师通过分析，给出了公司未来网络发展的如下需求。

(1) 部署交换机以太网，将用户终端的带宽提高至 100Mbps。

(2) 加强网络管理，提高网络安全，各个部门网络之间相互隔离。

(3) 适应公司网络规模的增长，包括计算机数量的增长和公司业务所需带宽的增长。

(4) 专门建设服务器群，为公司提供全局服务。

(5) 实现公司员工对各种 Internet 资源的访问。

(6) 设备模型有助于适应部门之间的数据流量的 20/80 规则。

(7) 提供数据访问的安全性和可靠性，支持多媒体应用。

(8) 对现有缆线进行升级，满足公司未来发展需要。

11.3.2 设计原则

根据上面对该公司目前网络使用情况以及公司网络将来发展的方向，网络设计采用如下的原则。

(1) 采用交换式以太网技术。

(2) 接入层设计：

① 为终端用户提供独占带宽。

② 将每个部门划分为独立的 VLAN。

③ 网络设备支持网络管理。

(3) 分布层/核心层设计：

① 采用三层交换技术，并提供 VLAN 间路由。

② 网络关键部分提供设备和链路冗余。

③ 千兆链路上行到核心层。

④ 服务器集中放置在靠近公司网络主干的位置。

在接入层，大量使用两层以太网交换机，为用户提供独占的宽带。根据公司各个部门主机的逻辑分布使用 VLAN 技术，使得部门的组织可以通过逻辑局域网来实现，专门服务各部门的服务器也划分到各自的 VLAN 内。网络中的设备都支持网络管理，提高网络的管理效率。鉴于该公司当前的规模，为了节约开支，可以考虑把分布层和核心层的功能部署在一台单独的设备中。如果公司规模进一步扩大，可以再将核心层和分布层的功能部署在不同的专用网络设备上。

在分布层/核心层，可以考虑使用三层交换机，一方面提供了快速的交换速率；另一方面也支持了 VLAN 的划分和 VLAN 间的通信。在分布层，可以考虑一些网络策略的使用，比如子网规划、地址汇总、路由协议、访问控制列表、远程访问等。服务器的集中放置，必然使得服务器网段数据量过大，因此为了保证足够的带宽，服务器应直接连接到分布层或核心层设备上，通过千兆以太网链路来连接。

11.3.3 设计方案一

图 11.9 所示是针对该公司当前网络的一个规划方案。在接入层选用 H3C S3600 交换机。分布层与核心层合二为一，选用了 H3C S7500 交换机，提供高速的数据交换。为了保证重要数据的高速访问以及链路冗余，采用了端口聚合技术。

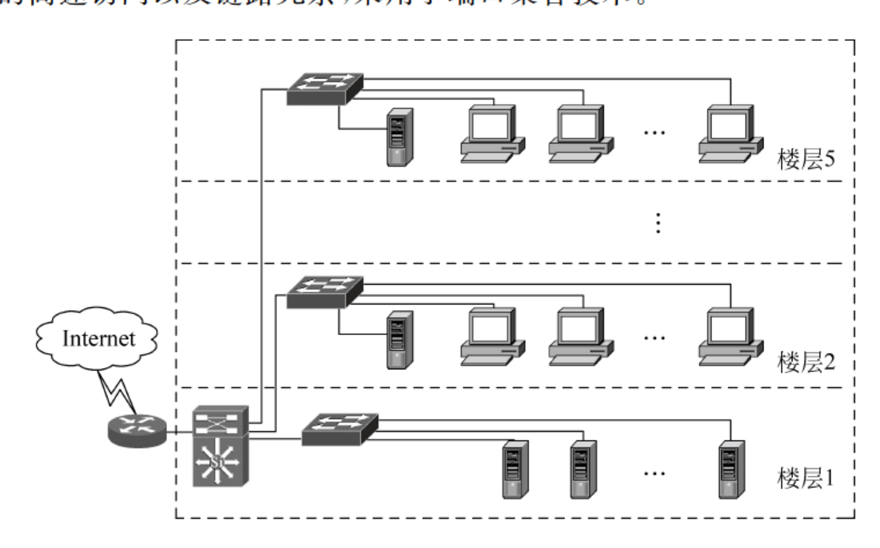


图 11.9 楼宇网络结构设计一

各楼层使用的水平缆线均重新敷设为 5e 类双绞线，整个水平子系统和工作区子系统所用的所有配件均达到 5e 标准；垂直子系统中敷设为 1000Base-SX 双模光纤，楼层中交换机也均安装了光纤模块用于连接；服务器群中的服务器均安装多网卡实现负载均衡。

这个方案也存在一些不足，首先是整个网络虽然通过交换机对网络进行了微分段(Microsegmentation)，即将整个网络划分为尽量多的冲突域，但整个网络仍然是在一个广播域下，广播包对网络的影响依然存在；其次是网络连接没有提供冗余，比如，分布层/核心层的 H3C S7500 交换机发生故障时，将导致楼层间不能访问。

为了提高网络的效率和健壮性、满足公司成长的需要，还可以考虑采用设计方案二。

11.3.4 设计方案二

方案二设计的网络结构模型层次划分明确，即接入层、分布层与核心层，如图11.10所示。

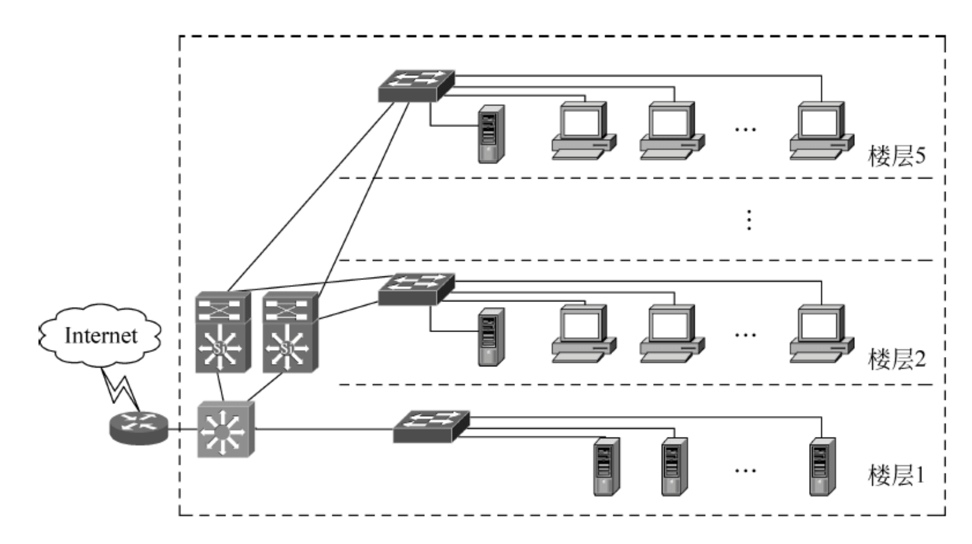


图11.10 楼宇网络结构设计二

接入层选用H3C S3600交换机(除了服务器网段外)，为主机提供100Base-T线缆接入；各层主机依据部门划分VLAN；服务器集中放置在楼层1的机房，各台服务器使用100Base-T连接到服务器群专用的H3C S3600交换机，保证服务器的带宽。

分布层选用两台华为Quidway S5500交换机，与接入层交换机分别互联，提供主备链路，保证可靠性。由于使用了VLAN技术，为实现逻辑的子网划分与通信，可以在H3C S5500交换机上配置路由策略，规划子网，配置路由协议，实施安全访问控制，配置QoS等。

核心层应该保证足够的宽带，快速数据传输。设备可以选用H3C S7500交换机，与分布层交换机和服务器网段交换机采用1000Base-FX光纤连接，与服务器相连接口采用链路聚合技术保证链路冗余。所有分布层设备、核心层设备以及服务器、NMS等均放置在机房，方便统一管理。

该方案大量采用了较昂贵的光纤传输介质、高性能交换机，以及为了提供设备和链路的冗余性，增加了设备和缆线的敷设，这样必然增加了该公司网络建设的成本。所以，最后在方案选择上根据用户的具体情况选择较合适的一种。

11.4 园区网络的设计案例

园区网络是指在园区内多个建筑物之间的网络连接，在园区网络中通常要求冗余可靠的连接。一个园区网络使用高带宽的物理层介质连接着相距较近的多个建筑物。通常物理层介质是由网络的所有者自己建设。

图 11.11 中使用三层交换机将园区中的建筑物连接起来。为了保证在园区网中分层设计方法得以执行,在网络层有效地控制广播,每个建筑物中分配连续的子网地址,以达到最相近的地址聚合。大型的园区网络方案应该能够支持高带宽的应用,如 VoIP、视频会议等。

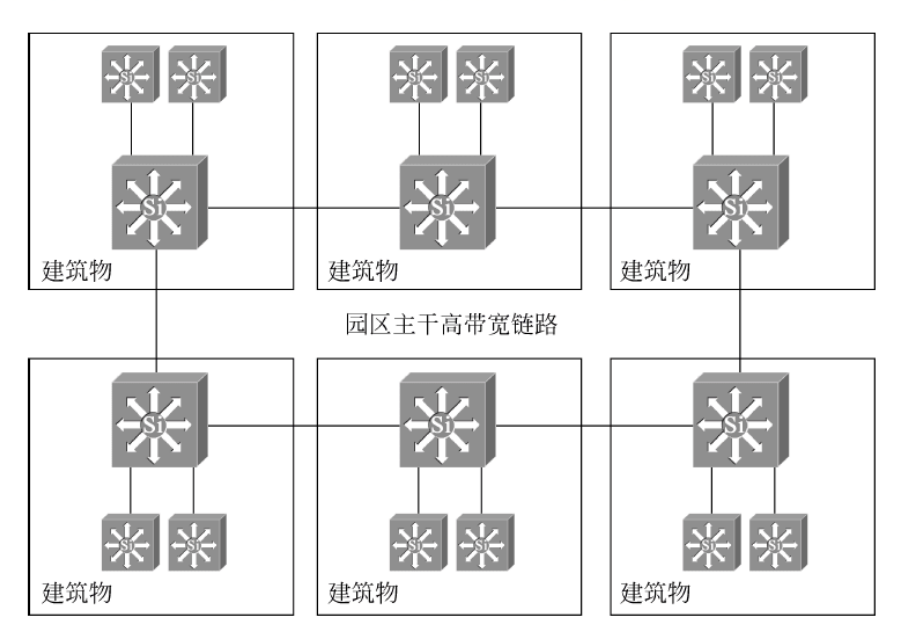


图 11.11 园区网络设计

某大学是一所地方大学,由 12 个二级学院及众多部门组成。该大学目前已经使用了网上教务管理系统及网上教学辅助系统等,并拥有许多各学科专业实验室与公共实验室,基本上每个实验室内都配有大量的计算机设备,但由于各个学院和部门先前都是根据自己的需要而组建的网络,采用了软硬件平台、网络拓扑结构都各不相同,因此随着学校网络的发展和整个学校的网络整合,现有的网络拓扑结构与性能也不能适应学校发展的需要,具体表现如下。

(1) 各学院自身的网络应用系统在组建时,往往只是从自身教学、科研的需要出发,因此存在大量的信息冗余和信息冲突。

(2) 各学院的各个网络终端使用自身的传输线路,传输速率和传输质量不等,不利于统一管理,随着终端数量的增多,此问题将日益突出。

(3) 由于之前的网络组建都是由本学院自己开发,因此存在多种机制、多种操作系统、多种协议、网络异构等情况。所以很难实现资源共享、系统互访、统一管理,网络系统的集成难度大大增加。

(4) 大部分系统没有设计为 C/S 模式或 B/S 模式,系统使用不方便。

(5) 校园内大多数计算机还不能充分利用学校拥有的 IP 地址资源访问 Internet,网络资源利用率低。

11.4.1 园区网络需求分析

由于现有应用系统存在上述缺陷,因此必须对现状和需求进行科学的分析,制定出全局网络的规划,既能满足发展,又要容纳现有系统,具体要求如下。

(1) 对目前各网络系统做大规模修改,各学院自己的网络系统应能平滑地过渡到整个

校园网中。

（2）提供各种灵活多变的联网方式，系统要有一定的可扩充性和可扩展性。

（3）提供高速平台和足够的带宽，为将来的 OA 系统、图像系统、远程教学、多媒体教学等应用提供一条可靠、健壮的"信息高速公路"。

（4）必须对整个校园网进行有效的集中管理。

（5）可以为校园内各个系统之间提供邮件服务、BBS 服务、文件服务、Web 服务等多种 Internet 服务。

（6）随着多媒体教学、远程教育、图像监控等行业的开展，校园网本身的业务规范将不断扩大，对学校网络的性能提出了新的要求。

11.4.2　设计原则

该园区网在设计时应遵循如下的系统设计原则。

（1）实用性：在满足网络建设需求的前提下，尽量做到网络规划达到够用就好，另外还要保护之前对网络的已有投资，充分利用现有的各种网络线路和网络设备，使得建成的园区网不仅实用、安全、可靠，易管理、易维护、易扩展，而且具有最高的性价比。

（2）开放性：网络系统的开放是网络技术发展的方向，因此在整个园区网的设计中采用的规范、技术要与设备供应商无关，具有较强的兼容性，便于与其他异构系统能够进行平滑的联动。

（3）先进性：当前的网络技术发展日新月异，方向把握不准则可能导致在很短时间内技术落伍，从而面临被淘汰的危险。因此在坚持实用性的前提下尽量采用国际先进的网络技术和设备，适合未来的发展，做到一次规划长期受益。

（4）可扩充性：所选择的联网方案及设备要能适应网络规模不断扩大的要求，以便于将来网络容量的扩充，并且适应信息技术的不断发展，平稳地向未来新技术过渡。

（5）可靠性：系统设计除采用信誉好、质量高的设备外，还应采用一系列容错、冗余技术，提高整个系统的健壮性。

（6）安全性：安全性包括信息的安全和网络的安全，这些都可以在园区网中通过部署在网络设备上的安全特征来获得支持。

11.4.3　园区网络设计方案一

1. 应用模式设计方案

对 ABC 大学的校园网这种规模大、集成度高的网络，建议采用 C/S 结构模式，即将网络结构建立在各类信息分布处理和集中管理相结合的方式上；由于将数据处理工作放在各终端独立处理，减轻了服务器的负担，设备的性能可以得到充分的发挥，而且信息资源可以分布共享、集中管理，使得系统的可靠性、开放性不单单依赖服务器，互补性很强。这种结构灵活性好，速度快，可靠性高，是当今流行的网络系统方案。

2. 网络带宽设计方案

采用交换机式以太网方案，交换式以太网是以常规以太网为基础的，它为每个节点提供

专用的以太网连接，为该段提供独占的宽带。这种方式对于那些要求专用服务的应用，如视频会议和其他自然数据行应用，是一种理想的选择，因为它确保应用保持较低时延。交换式以太网保留了传统以太网的帧格式，因此它可以保留现有以太网的基础设施。

用户接入层选用 H3C S3600 交换机。由于各台主机与交换机距离较近，交换机下行电缆选用 5 类双绞线；而各学院所在办公大楼较大，H3C S3600 交换机的距离已经超过了电缆线的标准 100m，因此选用 100Base-FX，为上行线路提供 100Mbps 宽带。

分布层设备汇聚了网络的大量流量，要求较高的数据转发速率，并且实施路由策略和安全控制。因此，分布层设备选用 H3C S7500 交换机。H3C S7500 交换机基于自适应安全网络的技术理念，在提供稳定、可靠、安全的高性能 L2/L3 层交换服务基础上，进一步提供了业务流分析、基于策略的 QOS、可控组播等智能的业务优化手段，从而为企业 IT 系统构建面向业务的网络平台，为实现通信整合、数据整合奠定了基础。

核心层主要功能是实现远程站点之间的优化传输，核心层设备要求很大的数据传输能力和路由负载平衡。选择 H3C S9500 交换机，该型号的交换机是面向以业务为核心的企业网络架构而推出的新一代核心路由交换机，该产品基于自适应安全网络的技术理念，在提供大容量、高性能 L2/L3 层交换服务基础上，进一步融合了硬件 IPv6、网络安全、网络业务分析等智能特性，可为企业构建融合业务的基础网络平台，进而帮助用户实现 IT 资源整合的需求。

3. 可靠性设计方案

由于校园网数据量大，访问人员多，同时承担了大量数据的传输、处理等工作，对分布层设备的容错性提出了很高的要求。为了增强网络的可靠性和可扩展性，分布层设备选用了两台 H3C S7500 交换机，两台交换机互为备份。接入层设备分别连接到这两台交换机上，实现冗余备份，确保网络的坚固性。同时，还需合理分配路由负载的平均，控制路由表的大小，实施路由策略。

4. 服务器管理方案

ABC 大学的 Web 服务器数量较多，并且每天的访问量很大，为了加强管理和维护，优化服务器数据传输速率，决定采用服务器集中放置的方式，并为服务器网段采用 1000Mbps 高速带宽。

5. 网络管理方案

网络管理系统应主要对网络资源的配置管理、故障管理、性能管理、记账管理、安全管理等方面有全面、有效的解决方案。

11.4.4 园区网络设计方案二

本校园网建设的总目标是：将学校各学院和各职能部门现有的及将来要配备的各种计算机、工作站和终端通过高性能的网络设备有效地相互连接起来，组成分布式、开放性的校园网网络环境，由网络主服务器进行统一管理和调控，以实现相互交流信息、共享网络资源和公共信息，使学校内各类信息资源更有效地为决策者、管理人员、教师和学生提供服务，并通过 CERNET 和 Internet 与国内外的其他校园网络相连，以方便与外界的学术及信息交流，提高学校的教学、科研和管理水平，实现校园内外的信息传输网络化、办公自动化、通信

现代化。

(1) 连接本校总部4个教学楼内的各学院和职能部门,以及分散于各城区的校区和办学点。

(2) 近期内可以支持近2000个独立用户。

(3) 实现资源共享。

(4) 提供丰富的网络服务。

(5) 提供100Mbps以上的高效的网络速率。

(6) 与CERNET和Internet连接,与国内外其他网络连接。

1. 用户需求分析

(1) 使学院的教学、科研和管理规范化、信息化:建立大学的校园网,可以使学校内部各部门之间相互及时交流信息,共享网络资源。

(2) 方便与外界的交流及沟通:使校园网能够实现国内、国际的信息传输,提供资源更多、范围更广的网络服务,是校园网建设的重要任务。

(3) 为其他入网用户提供服务:为分散办学点的教师、科研人员、校领导以及个人办公地点提供网络服务措施,使他们通过远程入网后能访问校园网、Internet网上的信息资源,进行国内、国际的网络通信。

2. 设计原则

(1) 结合本校现有计算机的现状和发展方向,兼容和包含现有设备,保护原有硬件和软件资源。

(2) 保证网络系统先进性、稳定性、可靠性、安全性和可维持性。

(3) 采用先进成熟的技术和设备,使建成的系统若干年内不落后。

(4) 坚持开放性,采用国际标准,以便今后系统扩充及升级。

(5) 坚持实用性原则。

3. 网络系统方案设计

1) 网络的节点设置及拓扑结构

从学院的实际情况出发,根据各学院内部的职能部门及教师办公室之间进行信息交换,学校行政机关各职能部门与各学院相应的职能部门之间进行信息交换,本网络采用星形拓扑结构。以计算中心的主服务器为中心,呈辐射状连接到其他各学院的子网,如图11.12所示。这种拓扑的最大特点就是当一条干线出现故障时,并不会影响其他干线的正常工作,所以可靠性相对较强。

2) 校园网主干网及各子网局域网网络选型

根据目前局域网技术的发展情况和实际需求,校园网的主干网选用千兆以太网,整个网络以1Gbps的交换机为中心,设立各类服务器,下设100Mbps交换器,连接各个子网和其他附属设备。

各学院的子网可选择100Base-TX,这是用于光纤和两对5类双绞线的快速以太网。工作站通过两对5类双绞线或光纤连接到100M的集线器或交换机上。

3) 广域网网络环境的选择

校园网在建成之后,需要接入CERNET,并通过它接入Internet。选择和本地最近的一个CERNET主节点之间的互联方式。

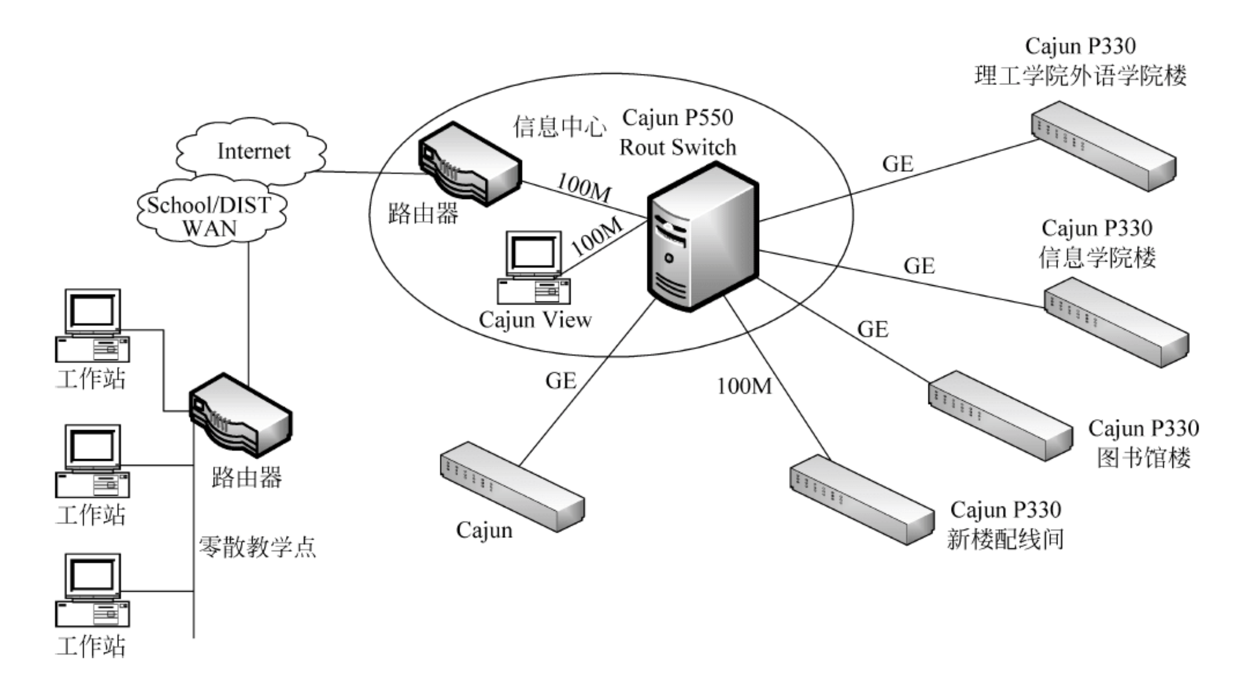


图 11.12　网络拓扑结构

由于学校总部地处闹市,周围建筑与道路众多,不适合自己在两者之间铺设光纤或进行微波通信方式,故选择租用电信部门的公用通信网。

根据对广域网网络资源的分析比较以及实际情况的需求,选择采用 DDN/Frame Relay 组网。因为它们网络伸缩性强,性价比高,适用于目前以及未来业务进一步发展的需求。

4) 网络互联方案

校总部的网络通过路由器和 DDN 专线实现与 CERNET 以及 Internet 的联通。其他校区及分散办学点根据自身规模及信息量的大小,目前可选择通过 ISDN 专线拨号接入 Internet,待以后网络规模扩大之后,再通过路由器和 DDN 专线接入 Internet。

由于距离太远,目前信息交换量不大,暂不考虑其他校区和分散办学点与校总部网络的直接连接。

5) 网络设备

主服务器是整个网络系统的核心,也是数据存储和处理的中心。因此,对它的可靠性、安全性和运算速度均有很高的要求。在这里选择 SUN 公司的企业级服务器。

网络主节点的核心设备选用一台朗讯公司的千兆位路由交换机 Cajun P550。它是具有路由功能的千兆位骨干交换机,既提供 1000Base-SX 的光纤模块,用于与各建筑物的中心交换机相连,又提供 10/100Base-TX 第二层交换模块,用于服务器的接入。

网络设备还包括两层交换机、网络适配器等。

6) 网络布线

楼内的布线采用 5 类非屏蔽双绞线。楼与楼之间采用光缆进行连接。光缆具有传输速率高、抗干扰能力强、可靠性高等优点,是建筑物间传输介质的首选。

各楼内采用结构化综合布线系统。各楼层安装配线架,用 5 类非屏蔽双绞线从各楼层的配线架布到各网络节点所在的墙壁。

从主节点交换机接出的线,通过光缆连接到各个楼的两层交换机上。两层交换机接出

的线，通过5类双绞线接到各楼层的Hub上，Hub与配线架相连，再通过楼内水平布线，通到各网络节点的墙壁接线盒上，再与计算机相连，实现数据传输。

7）网络安全措施

计算机网络系统的安全是网络设计中需要充分考虑的一个非常重要的问题。在设计中通常通过几个方面来保证网络系统及其数据的安全，主要包括：保证硬件可靠性；利用磁盘镜像技术，在数据库服务器中配置两个硬盘，设置容错方案；数据备份方案；用UPS电源以确保服务器及其硬盘的安全；设置网络防火墙；对用户访问服务器的权限进行限制等。

此外，整个网络还可以实现网端到网端的VLAN划分，在物理层上将用户分开，实现网络安全；可以将PC的MAC地址和IP地址在交换机上绑定，保证重要用户的安全性；可以通过IP地址或地址段(IP子网)来限制路由交换，实现网络安全。

4. 软件选择

(1) 网络操作系统的选择。在服务器上，可根据需要分别选择Linux或Windows 2000 Server。这两种操作系统适合于小型局域网，具有使用方便、简单的优点。

(2) 网络管理软件选择。

(3) 数据库选择。

5. 评价与经济可行性分析

1）系统处理能力

本网络以先进的快速网络交换技术为核心，采用性能优越的网络设备，确保主干网和主服务器的网络通道独享100Mbps的传输带宽，充分发挥服务器的强大的并行处理能力，为网上用户提供及时、高效的网络服务。所有用户都分别纳入相应的以太网段，保证网络所有节点都能享用足够的网络带宽，保证网络的通畅。整套系统采用先进的客户机/服务器结构。信息处理任务可以由主服务器、分服务器以及客房机共同承担，所以系统的处理能力能够很好地满足该系统的业务处理需求。

采用交换式以太网作为网络主干，使整个网络具有先进、快速、实用、灵活和经济的特点，符合网络技术发展潮流，比传统的共享式以太网或FDDI在技术上更先进，网络效率更高，还大大降低了投资成本，并为向ATM网络发展奠定了基础。

采用网络交换机能大幅度提高网络性能，不仅体现在交换机能为服务器和网络各节点提供各自所需要的网络带宽，缓解网络拥挤情况，而且交换机允许不同的网络节点同时发送数据包，克服了传统以太网的局限。

2）系统可靠性

先进的网络拓扑结构和结构化综合布线系统，能够最大限度地提高网络的可靠性，如果某一点出现故障，不会影响整个网络的运行，并且便于维护。

网络交换机能够对数据包进行过滤，起到一定程度的防火墙作用。

网络管理软件能对整个网络实施有效的控制和管理，并进行故障诊断和排除。

主服务器本身具有自动系统恢复功能，操作系统具有C2级的安全保密系统。主要的网络设备都是高可靠性产品，整个系统可靠性很高。

3）系统扩充性

主交换机有冗余插槽，可供以后提供更多的网络交换端口和安插ATM模块。

服务器本身具有一倍以上的可扩充能力，以后在出现数据量大幅度增加的情况下，只需

通过增加服务器的系统板和 CPU 模块，就可以使服务器的处理能力提高。

主要网络设备均具有 ATM 扩充能力，在需要建立 ATM 网络时，能实现连接。

通过路由器，本系统与其他网络的连接就变得很简单。根据需要，可以通过 X.25、DDN 或帧中继与远程节点建立连接。

4）经济可行性

本系统采用客户机/服务器结构，因此整个系统不仅处理能力很强，而且价格比传统的大型机系统也要便宜许多，而且以后的扩充费用也比较低。主服务器和中心交换机产品不仅性能比较优越，其每个端口的性价比也名列前茅。采用系统集成方式，各种设备实现了最优组合和搭配，在充分考虑系统冗余和以后扩充的前提下，尽量做到设备的物尽其用，实现最佳的性价比。

课后习题

1. 网络工程集成设计的步骤和原则有哪些？
2. 网络布线原则有哪些？
3. 局域网系统设计的主要内容有哪些？
4. 广域网系统设计的主要内容有哪些？
5. 园区网络通常的体系结构是什么？每一层的主要功能是什么？

参 考 文 献

[1] W Richard Stevens. TCP/IP 详解卷 1：协议[M]. 范建华，等，译. 北京：机械工业出版社，2000.
[2] 谢希仁. 计算机网络[M]. 7 版. 北京：电子工业出版社，2017.
[3] 雷震甲. 网络工程师教程[M]. 2 版. 北京：清华大学出版社，2006.
[4] Andrew S Tanenbaum. 计算机网络[M]. 5 版. 潘爱民，译. 北京：清华大学出版社，2012.
[5] James F Kurose. 计算机网络——自顶向下方法与 Internet 特色[M]. 影印版. 北京：高等教育出版社，2001.
[6] Walter Goralski. 现代 TCP/IP 网络详解[M]. 北京：电子工业出版社，2015.
[7] 竹下隆史，等. 图解 TCP/IP[M]. 5 版. 北京：人民邮电出版社，2013.
[8] 曲大成，等. Internet 技术与应用教程[M]. 北京：高等教育出版社，2007.
[9] 王晓军. 深入理解计算机系统[M]. 北京：机械工业出版社，2016.
[10] 高传善，曹袖，等. 计算机网络[M]. 北京：高等教育出版社，2013.
[11] Douglas E Comer. 用 TCP/IP 进行网际互联(第一卷)[M]. 4 版. 林遥，等，译. 北京：电子工业出版社，2001.
[12] 陈鸣. 计算机网络：原理与实践[M]. 北京：机械工业出版社，2013.
[13] 韩立刚，等. 计算机网络原理创新教程. 北京：中国水利水电出版社，2017.
[14] 王达，等. 深入理解计算机网络[M]. 北京：机械工业出版社，2013.